TRANSACTIONS
OF THE
AMERICAN INSTITUTE
OF
ELECTRICAL ENGINEERS

VOLUME 76
1957

PART II. APPLICATIONS AND INDUSTRY

PUBLISHED BY THE
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
33 WEST 39TH STREET
NEW YORK 18, N. Y., U. S. A.

Copyright 1958 by the American Institute of Electrical Engineers
Printed in U. S. A. by The Mack Printing Company, Easton, Pa.

Preface

The AIEE *Transactions* for 1957 (volume 76) is published in three parts: Part I. Communication and Electronics; Part II. Applications and Industry; and Part III. Power Apparatus and Systems. The papers in each of the three parts are classified according to subject matter as follows:

Part I. Communication and Electronics

- Communication Switching Systems
- Communication Theory
- Radio Communications Systems
- Special Communications Applications
- Telegraph Systems
- Television and Aural Broadcasting Systems
- Wire Communications Systems
- Basic Sciences
- Computing Devices
- Electrical Techniques in Medicine and Biology
- Electronics
- Instruments and Measurements
- Magnetic Amplifiers
- Metallic Rectifiers
- Nucleonics
- Solid State Devices

Part II. Applications and Industry

- Air Transportation
- Domestic and Commercial Applications
- Land Transportation
- Marine Transportation
- Production and Application of Light
- Chemical Industry
- Electric Heating
- Electric Welding
- Feedback Control Systems
- General Industry Applications
- Industrial Control
- Industrial Power Rectifiers
- Industrial Power Systems
- Mining and Metal Industry
- Petroleum Industry

Part III. Power Apparatus and Systems

- Carrier Current
- Dielectrics
- Insulated Conductors
- Power Generation
- Protective Devices
- Relays
- Rotating Machinery
- Substations
- Switchgear
- System Engineering
- Transformers
- Transmission and Distribution

- Education
- Safety
- Research

Each part has been indexed separately in the back of that particular part. The three parts are not cross-referenced; hence the user should determine first whether the subject matter of the paper desired is in the field of communication and electronics, applications and industry, or power apparatus and systems. Papers are listed in the subject index under several key words in the titles. The original numbers assigned to the papers are given in the author index. Volume 76 contains the technical papers and related discussions presented at these meetings:

1. Winter General Meeting, New York, N. Y., January 21–25, 1957.
2. Southern District Meeting, Jackson, Miss., April 3–5, 1957.
3. Great Lakes District Meeting, Des Moines, Iowa, April 15–17, 1957.
4. North Eastern District Meeting, Pittsfield, Mass., May 1–3, 1957.
5. Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957.
6. Pacific General Meeting, Pasco, Wash., August 28–30, 1957.
7. Fall General Meeting, Chicago, Ill., October 7–11, 1957.

Statements and opinions given in the papers and discussions published in *Transactions* are the expressions of the contributors for which the American Institute of Electrical Engineers assumes no responsibility.

Applications and Industry Index for 1957

1. Technical Subject Index

A-C Aircraft Electric Power Systems, A Method of Fault Analysis for. Rindt, Jessee..........338-44
A-C Generator, A Brushless Air-Cooled Aircraft. Smith ....189-91; disc. 191
A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips, Protecting. Brightman, McGee, Reifschneider..........114-19
Adaptive Servomechanisms. Drenick, Shahbender..........286-91; disc. 291
Air-Borne Transformers, Size Reduction of. Lee..372-6
Air Conditioning, Single-Phase Versus Three-Phase Service for Residential. Anderson, Hutchinson..........61-4; disc. 64
Air-Cooled Aircraft A-C Generator, A Brushless. Smith ....189-91; disc. 191
Air Cooling Systems for Aircraft Generators, Ram-. Moroney..........217-21
Air Force Weapons, Ground-Power Equipment for the Support of. Holliday..........237-43
(Air Transportation) Electric Ignition for Conductive Fuels. Sheheen..........233-6
(Air Transportation) Flat Work-Coil Design. Mathews. 249-56
(Air Transportation) Informative Feedback in Jet-Pilot Control-Stick Motion. Diamantides..........243-9
(Air Transportation) Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines. Robinson..........192-9
(Air Transportation) Salient-Pole Permanent Magnet Alternators for High Speed Drive. Puder, Strauss..........333-8
(Air Transportation) Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator. Robinson, Tse..........199-204
(Air Transportation) The Role of Analog Computers in Propeller Control Design. Frick..257-63; disc 263
Aircraft D-C Power Systems with Inverter Lodas, The Stability of. Thomas..........183-8
Aircraft, Design of Generating Equipment for Electra Turboprop. Alger, Warden, Hansen, Klein..........320-32
Aircraft Electric Power Systems, A Method of Fault Analysis for A-C. Rindt, Jessee..........338-44
Aircraft Electric Power Systems, Static Control and Protection of. Hulsey, Schuh..........209-13
Aircraft Electric Systems, Development of Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps. Markowitz.222-32
Aircraft Generator Protective Systems, Some Applications of Magnetic Amplifiers in. Plette, Butler..........427-33
Aircraft Power Systems, Voltage Modulation on. Klokow, Yohe..........314-19
Aircraft, Silver-Zinc Batteries as Source Primary Electric Power for Pilotless. Abrahamson.....297-300
Aircraft, Thermal Consideration of Generators in High-Speed. Kusko..........205-08
Alternator, Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled. Robinson, Tse....199-204
Alternators for High Speed Drive, Salient-Pole Permanent Magnet. Puder, Strauss..........333-8
American Railroads, The Case for a Standard System of Railroad Electrification for the. Perkinson..........101-04; disc. 104
Amplidyne Servo Network, Transfer Matrix Stability Criterion Applied to an. Honnell, Wolffenstein..........143-8; disc. 148
Amplifiers, A Study of the Transfer Function of Contact-Modulated. Krantz, Salati, Berkowitz.......23-8
Amplifiers with Feedback, The Response of Relay. Gibson, Tuteur..........303-07
Analog Computers in Propeller Control Design, The Role of. Frick..........257-63; disc. 263
Analysis for A-C Aircraft Electric Power Systems, A Method of Fault. Rindt, Jessee..........338-44
Analysis of Irregular Current and Voltage Wave Shapes Found in Industrial Power Engineering. Kaufmann..........1-5
Analytic Method for Finding the Closed-Loop Frequency Response of Nonlinear Feedback Control Systems, An. Ogata..........277-85
Appliances, Some Problems and Design Factors Concerning Noise in Motor-Driven Home. Cunningham, Wolfert..........17-22
Application of a Mechanical Rectifier to a Cyclotron. Warren..........273-6; disc. 277
Application of D-C Machines to Oil-Well Drilling. Hefner..........345-50; disc. 350
Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles, Criteria for Industrial. Munson, Germain..........379-82
Arc-Furnace Corrective Equipment Using a High Value of Buffer Reactance. Harper, Macon, Sedgwick..........74-9; disc. 79
Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker, Survey of. (Committee Report)..........170-81; disc. 181
Arc-Furnace Loads, Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying. Concordia, Levoy, Thomas.123-34; disc. 134
Automatic Power Plants, A 10-Kw Germanium Rectifier for. Hake..........361-6
Automotive Vehicles, Selenium Rectifier Applications in. Szabo..........369-71
Auxiliaries, The Lightweight Train—Its Power Supply for. Swarner..........409-15; disc. 415

B

Batteries as Source Primary Electric Power for Pilotless Aircraft, Silver-Zinc. Abrahamson.......297-300
Blast-Cooled Alternator, Surface Heat-Transfer Coefficients of Salient Poles in a. Robinson, Tse..199-204
Brushless Air-Cooled Aircraft A-C Generator, A. Smith ....189-91; disc. 191
Buffer Reactance, Arc-Furnace Corrective Equipment Using a High Value of. Harper, Macon, Sedgwick..........74-9; disc. 79
Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads, Selection of. Concordia, Levoy, Thomas......123-34; disc. 134

C

Cables Installed in the Same Residence for Data Purposes, Heat Pump and Heating. Jones, Linden..........307-14
Calculation of Induction-Heating Coils, Design and. Baker..........31-8; disc. 39
Case for a Standard System of Railroad Electrification for the American Railroads, The. Perkinson..........101-04; disc. 104
Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control Valves, Generalized. Zaborszky, Harrington..........401-08; disc. 408
(Chemical Industry) Direct Current Metering of Large Electrochemical Pot-Lines. Reagan.......300-02
(Chemical Industry) Silver-Zinc Batteries As Source-Primary Electric Power for Pilotless Aircraft. Abrahamson..........297-300
Circuit Breaker Series Trips, Protecting A-C Motors with Low-Voltage Air-. Brightman, McGee, Reifschneider..........114-19
Closed-Loop Frequency Response of Nonlinear Feedback Control Systems, An Analytic Method for Finding the. Ogata..........277-85
Closed-Loop Poles of Control Systems, A Quick Method for Estimating. Chen..........80-6; disc. 86
Coefficients of Salient Poles in a Blast-Cooled Alternator, Surface Heat-Transfer. Robinson, Tse....199-204
Coil Design, Flat Work-. Mathews..........249-56
Coils, Design and Calculation of Induction-Heating. Baker..........31-8; disc. 39
(Committee Report) Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker. (57-9)..........170-81; disc. 181
Compensation of an Error-Sampled System by a Multirate Controller. Kranc.......149-58; disc. 158
Computer Analysis for the Design of Steel Mill Control Systems, Modern. Reider, Spergel..........105-09; disc. 109
Computers in Propeller Control Design, The Role of Analog. Frick..........257-63; disc. 263
Condensers on Power Systems Supplying Arc-Furnace Loads, Selection of Buffer Reactors and Synchronous. Concordia, Levoy, Thomas.123-34; disc. 134
Conductive Fuels, Electric Ignition for. Sheheen..........233-6
Contact-Modulated Amplifiers, A Study of the Transfer Function of. Krantz, Salati, Berkowitz....23-8
Contactor for Direct-Press Drives, An Ignitron. Pettit, Montross..........110-13; disc. 113
Contactor Servomechanisms, A New Method for Compensating. Embler, Weaver..........165-70
Control-Stick Motion, Informative Feedback in Jet-Pilot. Diamantides..........243-9
Control Systems, A Quick Metnod for Estimating Closed-Loop Poles of. Chen..........80-6; disc. 86
Control Systems, A Synthesis Method for Multipole. Freeman..........28-31
Control Systems, Modern Computer Analysis for the Design of Steel Mill. Reider, Spergel..........105-09; disc. 109
Control Valves. A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic. Zaborszky, Harrington..........394-401; disc 408
Control Valves, Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic. Zaborszky, Harrington..........401-08; disc. 408
Controller, Compensation of an Error-Sampled System by a Multirate. Kranc..........149-58; disc. 158
Controls in Linear Multivariable Systems, Noninteracting. Kavanagh..........95-9; disc. 99
Cooled Aircraft A-C Generator, A Brushless Air-. Smith..........189-91; disc. 191
Cooling Systems for Aircraft Generators, Ram-Air. Moroney..........217-21
Corrective Equipment Using a High Value of Buffer Reactance, Arc-Furnace. Harper, Macon, Sedgwick..........74-9; disc. 79
Criteria for Industrial Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles. Munson, Germain..........379-82
Criteria, The Selection and Use of Servo Performance. Schultz, Rideout..........383-7; disc. 387
Criterion Applied to an Amplidyne Servo Network, Transfer Matrix Stability. Honnell, Wolffenstein..........143-8; disc. 148
Criterion for Second-Order Saturated Servomechanisms, Extended Switching. Diesel..........388-93
Criterion of Linear Servomechanisms by a Graphical Method, A New Stability. Numakura, Miura..40-8
Current and Voltage Wave Shapes Found in Industrial Power Engineering, Analysis of Irregular. Kaufmann..........1-5
Cyclotron, Application of a Mechanical Rectifier to a. Warren..........273-6; disc. 277

D

Damping, Quasi-Optimization of Relay Servos by Use of Discontinuous. Harris, McDonald, Thaler..........292-6; disc. 296
Data Purposes, Heat Pump and Heating Cables Installed in the Same Residence for. Jones, Linden..307-14
D-C Machines to Oil-Well Drilling, Application of. Hefner..........345-50; disc. 350
D-C Power Systems with Inverter Loads, The Stability of Aircraft. Thomas..........183-8
Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems, Development of. Markowitz..........222-32
Density Selenium Rectifier, A High-Current-. Geib, Brown..........358-60
Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves, A. Zaborszky, Harrington.....394-401; disc. 408
Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive. McElhenny..........10-16
Design and Calculation of Induction-Heating Coils. Baker..........31-8; disc. 39
Design, Flat Work-Coil. Mathews..........249-56
Design of Generating Equipment for Electra Turboprop Aircraft. Alger, Warden, Hansen, Klein...320-32
Design of Ignitron Tubes for Rectifier Locomotives. Belnap, Zehner, Zuvers..........214-17
Design, The Role of Analog Computers in Propeller Control. Frick..........257-63; disc. 263
Development of Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems. Markowitz..........222-32
Diesel Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger Trains. Steinbrook..........136-42

Digital Flying Extensometer for Temper Rolling Mills, A. Wells.......376–8
Digital Telemetering and Print Out of Pipe Line Information. Nigh.......119–22; disc. 122
Direct Current Metering of Large Electrochemical Pot-Lines. Reagan.......300–02
Direct-Press Drives, An Ignitron Contactor for. Pettit, Montross.......110–13; disc. 113
Distribution System, A Transient Stability and Voltage Study for a Modern Oil Refinery. Boice, Durand, Dalasta.......159–64
Drilling, Application of D-C Machines to Oil-Well. Hefner.......345–50; disc. 350
Drive, Salient-Pole Permanent-Magnet Alternators for High-Speed. Puder, Strauss.......333–8
Drives, An Ignitron Contactor for Direct-Press. Pettit, Montross.......110–13; disc. 113

E

Economic Aspects of the Residential Heat Pump on Electric Utility Systems, Load and. Bary.......49–53; disc. 53
Economic Evaluation of Industrial Power-System Reliability. Dickinson.......264–9; disc. 269
8,500-Hp Gas-Turbine-Electric Locomotive, An. Smith.......5–10
Electra Turboprop Aircraft, Design of Generating Equipment for. Alger, Warden, Hansen, Klein...320–32
(Electric Heating) Design and Calculation of Induction-Heating Coils. Baker.......31–8; disc. 39
Electric Ignition for Conductive Fuels. Sheheen..233–6
Electric Locomotive, An 8,500-Hp Gas-Turbine-. Smith.......5–10
Electric Power for Pilotless Aircraft, Silver-Zinc Batteries as Source Primary. Abrahamson.......297–300
Electric Power Systems, A Method of Fault Analysis for A-C Aircraft. Rindt, Jessee.......338–44
Electric Power Systems, Static Control and Protection of Aircraft. Hulsey, Schuh.......209–13
Electric System of the Rectifier-Type Locomotives for the Virginian Railway, The. Brown.......68–73
Electric Utility Systems, Load and Economic Aspects of the Residential Heat Pump on. Bary.......49–53; disc. 53
Electrification for the American Railroads, The Case for a Standard System of Railroad. Perkinson.......101–04; disc. 104
Electrochemical Pot-Lines, Direct Current Metering of Large. Reagan.......300–02
Electrohydraulic Control Valves, A Describing Function for the Multiple Nonlinearities Present in 2-Stage. Zaborszky, Harrington.......394–401; disc. 408
Electrohydraulic Control Valves, Generalized Charts of the Effects in Nonlinearities in 2-Stage. Zaborszky, Harrington.......401–08; disc. 408
Electrolytic Processes, Germanium Rectifier Equipment for. Miller, Hodgson.......353–7
Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines, Prediction of Transient and. Robinson.......192–9
Equipment for Electra Turboprop Aircraft, Design of Generating. Alger, Warden, Hansen, Klein.320–32
Equipment for Electrolytic Processes, Germanium Rectifier. Miller, Hodgson.......353–7
Equipment for the Support of Air Force Weapons, Ground-Power. Holliday.......237–43
Equipment Using a High Value of Buffer Reactance, Arc-Furnace Corrective. Harper, Macon, Sedgwick.......74–9; disc. 79
Equipments, Silicon Power Rectifier. Marshall, Stewart.......366–9
Error-Sampled System by a Multirate Controller, Compensation of an. Kranc.....149–58; disc. 158
Evaluation of Feedback-Control Systems Subjected to Large Signals. Haas.......88–95
Evaluation of Industrial Power-System Reliability, Economic. Dickinson.......264–9; disc. 269
Extended Switching Criterion for Second-Order Saturated Servomechanisms. Diesel.......388–93
Extensometer for Temper Rolling Mills, A Digital Flying. Wells.......376–8

F

Fault Analysis for A-C Aircraft Electric Power Systems, A Method of. Rindt, Jessee.......338–44
(Feedback Control) A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves. Zaborszky, Harrington.......394–401; disc. 408
(Feedback Control) A New Stability Criterion of Linear Servomechanisms by a Graphical Method. Numakura, Miura.......40–8
(Feedback Control) A Quick Method for Estimating Closed-Loop Poles of Control Systems. Chen.......80–6; disc. 86
(Feedback Control) A Study of the Transfer Function of Contact-Modulated Amplifiers. Krantz, Salati, Berkowitz.......23–8
(Feedback Control) A Synthesis Method for Multipole Control Systems. Freeman.......28–31
(Feedback Control) Adaptive Servomechanisms. Drenick, Shahbender.......286–91; disc. 291
(Feedback Control) Compensation of an Error-Sampled System by a Multirate Controller. Kranc.......149–58; disc. 158
(Feedback Control) Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control Valves. Zaborszky, Harrington.......401–08; disc. 408
(Feedback Control) Noninteracting Controls in Linear Multivariable Systems. Kavanagh..95–9; disc. 99
Feedback Control System, A 2-Dimensional. Sarachik, Ragazzini.......55–60; disc. 60
Feedback Control Systems, An Analytic Method for Finding the Closed Loop Frequency Response of Nonlinear. Ogata.......277–85
Feedback-Control Systems Subjected to Large Signals, Evaluation of. Haas.......88–95
(Feedback Control) The Selection and Use of Servo Performance Criteria. Schultz, Rideout.......383–7; disc. 387
(Feedback Control) Transfer Matrix Stability Criterion Applied to an Amplidyne Servo Network. Honnell, Wolffenstein.......143–8; disc. 148
Feedback in Jet-Pilot Control-Stick Motion, Informative. Diamantides.......243–9
Feedback, The Response of Relay Amplifiers with. Gibson, Tuteur.......303–07
Flat Work-Coil Design. Mathews.......249–56
Flicker, Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp. (Committee Report).......170–81; disc. 181
Flying Extensometer for Temper Rolling Mills, A Digital. Wells.......376–8
400 to 20,000 Cycles, Criteria for Industrial Application of Single-Phase Transmission Lines. Munson, Germain.......379–82
Frequency Response of Nonlinear Feedback Control Systems, An Analytic Method for Finding the Closed Loop. Ogata.......277–85
Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems, Development of Definitions and Limits for. Markowitz.......222–32
Fuels, Electric Ignition for Conductive. Sheheen.......233–6
Function for the Multiple Nonlinearities Present in 2-State Electrohydraulic Control Valves, A Describing. Zaborszky, Harrington....394–401; disc. 408
Function of Contact-Modulated Amplifiers, A Study of the Transfer. Krantz, Salati, Berkowitz......23–8
Furnace Corrective Equipment Using a High Value of Buffer Reactance, Arc-. Harper, Macon, Sedgwick.......74–9; disc. 79
Furnace Installations on Power Systems and Resulting Lamp Flicker, Survey of Arc. (Committee Report).......170–81; disc. 181

G

Gas-Turbine-Electric Locomotive, An 8,500-Hp. Smith.......5–10
Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control Valves. Zaborszky, Harrington.......401–08; disc. 408
Generating Equipment for Electra Turboprop Aircraft, Design of. Alger, Warden, Hansen, Klein..320–32
Generator, A Brushless Air-Cooled Aircraft A-C. Smith.......189–91; disc. 191
Generator Protective Systems, Some Applications of Magnetic Amplifiers in Aircraft. Plette, Butler.......427–33
Generators in High-Speed Aircraft, Thermal Consideration of. Kusko.......205–08
Generators, Ram-Air Cooling Systems for Aircraft. Moroney.......217–21
Germanium Rectifier Equipment for Electrolytic Processes. Miller, Hodgson.......353–7
Germanium Rectifier for Automatic Power Plants, A 10-Kw. Hake.......361–6
Graphical Method, A New Stability Criterion of Linear Servomechanisms by a. Numakura, Miura.......40–8
Ground-Power Equipment for the Support of Air Force Weapons. Holliday.......237–43

H

Heat Pump and Heating Cables Installed in the Same Residence for Data Purposes. Jones, Linden.......307–14
Heat Pump on Electric Utility Systems, Load and Economic Aspects of the Residential. Bary.......49–53; disc. 53
Heat-Transfer Cofficients of Salient Poles in a Blast-Cooled Alternator, Surface. Robinson, Tse.......199–204
High-Current-Density Selenium Rectifier, A. Geib, Brown.......358–60
High-Speed Aircraft, Thermal Consideration of Generators in. Kusko.......205–08
High Speed Drive, Salient-Pole Permanent-Magnet Alternators for. Puder, Strauss.......333–8
Home Appliances, Some Problems and Design Factors Concerning Noise in Motor-Driven. Cunningham, Wolfert.......17–22
Hydraulic Transmission for Lightweight Passenger Trains, Diesel Locomotives with Mechydro. Steinbrook.......136–42

I

Ignition for Conductive Fuels, Electric. Sheheen..233–6
Ignitron Contactor for Direct-Press Drives, An. Pettit, Montross.......110–13; disc. 113
Ignitron Tubes for Rectifier Locomotives, Design of. Belnap, Zehner, Zuvers.......214–17
Induction-Heating Coils, Design and Calculation of. Baker.......31–8; disc. 39
Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad, Multiple-Unit-Rectifier Motive Power—. Hibbard, Garry, Loomis.......416–25; disc. 425
(Industrial) Application of a Mechanical Rectifier to a Cyclotron. Warren.......273–6; disc. 277
Industrial Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles, Criteria for. Munson, Germain.......379–82
(Industrial Control) A Digital Flying Extensometer for Temper Rolling Mills. Wells.......376–8
(Industrial Control) An Ignitron Contactor for Direct-Press Drives. Pettit, Montross...110–13; disc. 113
(Industrial Control) Protecting A-C Motors with Low-Voltage Air-Circuit Breaker Series Trips. Brightman, McGee, Reifschneider.......114–19
Industrial Power Engineering, Analysis of Irregular Current and Voltage Wave Shapes Found in. Kaufmann.......1–5
(Industrial Power Rectifiers) A High-Current-Density Selenium Rectifier. Geib, Brown.......358–60
(Industrial Power Rectifiers) A 10-Kw Germanium Rectifier for Automatic Power Plants. Hake.361–6
(Industrial Power Rectifiers) Germanium Rectifier Equipment for Electrolytic Processes. Miller, Hodgson.......353–7
(Industrial Power Rectifiers) Selenium Rectifier Applications in Automotive Vehicles. Szabo......369–71
(Industrial Power Rectifiers) Silicon Power Rectifier Equipments. Marshall, Stewart.......366–9
Industrial Power-System Reliability, Economic Evaluation of. Dickinson.......264–9; disc. 269
Informative Feedback in Jet-Pilot Control-Stick Motion. Diamantides.......243–9
Insulation Damage of Blast-Cooled Machines, Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative. Robinson.......192–9
Inverter Loads, The Stability of Aircraft D-C Power Systems with. Thomas.......183–8
Irregular Current and Voltage Wave Shapes Found in Industrial Power Engineering, Analysis of. Kaufmann.......1–5

J

Jet-Pilot Control-Stick Motion, Informative Feedback in. Diamantides.......243–9

L

Lamp Flicker, Survey of Arc-Furnace Installations on Power Systems and Resulting. (Committee Report).......170–81; disc. 181
(Land Transportation) An 8,500-Hp Gas-Turbine-Electric Locomotive. Smith.......5–10
(Land Transportation) Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive. McElhenny.......10–16
(Land Transportation) Diesel Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger Trains. Steinbrook.......136–42
(Land Transportation) Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad. Hibbard, Garry, Loomis.......416–25; disc. 425
(Land Transportation) The Case for a Standard System of Railroad Electrification for the American Railroads. Perkinson.......101–04; disc. 104
(Land Transportation) The Electric System of the Rectifier-Type Locomotives for the Virginian Railway. Brown.......68–73
(Land Transportation) The Lightweight Train—Its Power Supply for Auxiliaries. Swarner.......409–15; disc. 415
Lightweight Train—Its Power Supply for Auxiliaries, The. Swarner.......409–15; disc. 415
Linear Multivariable Systems, Noninteracting Controls in. Kavanagh.......95–9; disc. 99

Linear Servomechanisms by a Graphical Method, A New Stability Criterion of. Numakura, Miura........40–8
Load and Economic Aspects of the Residential Heat Pump on Electric Utility Systems. Bary........49–53; disc. 53
Loads, The Stability of Aircraft D-C Power Systems with Inverter. Thomas........183–8
Locomotive, An 8,500-Hp Gas-Turbine Electric. Smith........5–10
Locomotive, Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant. McElhenny........10–16
Locomotives, Design of Ignitron Tubes for Rectifier. Belnap, Zehner, Zuvers........214–17
Locomotives for the Virginian Railway, The Electric System of the Rectifier-Type. Brown........68–73
Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger Trains, Diesel. Steinbrook........136–42
Loop Poles of Control Systems, A Quick Method for Estimating Closed-. Chen........80–6; disc. 86
Low-Voltage Air-Circuit-Breaker Series Trips, Protecting A-C Motors with. Brightman, McGee, Reifschneider........114–19

M

Machines, Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled. Robinson........192–9
Machines to Oil-Well Drilling, Application of D-C. Hefner........345–50; disc. 350
Magnet Alternators for High Speed Drive, Salient-Pole Permanent. Puder, Strauss........333–8
Magnetic Amplifiers in Aircraft Generator Protective Systems, Some Applications of. Plette, Butler........427–33
Matrix Stability Criterion Applied to an Amplidyne Servo Network, Transfer. Honnell, Wolffenstein........143–8; disc. 148
Mechanical Rectifier to a Cyclotron, Application of a. Warren........273–6; disc. 277
Mechydro Hydraulic Transmission for Lightweight Passenger Trains, Diesel Locomotives with. Steinbrook........136–42
Metering of Large Electrochemical Pot-Lines, Direct Current. Reagan........300–02
Method for Compensating Contactor Servomechanisms, A New. Embler, Weaver........165–70
Method for Finding the Closed-Loop Frequency Response of Nonlinear Feedback Control Systems, An Analytic. Ogata........277–85
Method for Multipole Control Systems, A Synthesis. Freeman........28–31
Method of Estimating Closed-Loop Poles of Control Systems, A Quick. Chen........80–6; disc. 86
Method of Fault Analysis for A-C Aircraft Electric Power Systems, A. Rindt, Jessee........338–44
Modern Computer Analysis for the Design of Steel Mill Control Systems. Reider, Spergel........105–09; disc. 109
Modulated Amplifiers, A Study of the Transfer Function of Contact-. Krantz, Salati, Berkowitz........23–8
Modulation on Aircraft Power Systems, Voltage. Klokow, Yohe........314–19
Motion, Informative Feedback In Jet-Pilot Control-Stick. Diamantides........243–9
Motive Power—Inductive Co-ordination Considerations on the New York, New Haven, and Hartford Railroad, Multiple-Unit-Rectifier. Hibbard, Garry, Loomis........416–25; disc. 425
Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive, Design and Analysis of the. McElhenny........10–16
Motor-Drive Home Appliances, Some Problems and Design Factors Concerning Noise in. Cunningham, Wolfert........17–22
Motors with Low-Voltage Air-Circuit Breaker Series Trips, Protecting A-C. Brightman, McGee, Reifschneider........114–19
Multiple Nonlinearities Present In 2-Stage Electrohydraulic Control Valves, A Describing Function for the. Zaborszky, Harrington...394–401; disc. 408
Multiple-Unit-Rectifier Motive Power—Inductive Coordination Considerations on the New York, New Haven and Hartford Railroad. Hibbard, Garry, Loomis........416–25; disc. 425
Multipole Control Systems, A Synthesis Method for. Freeman........28–31
Multirate Controller, Compensation of an Error-Sampled System by a. Kranc...149–58; disc. 158
Multivariable Systems, Noninteracting Controls in Linear. Kavanagh........95–9; disc. 99

N

Network, Transfer Matrix Stability Criterion Applied to an Amplidyne Servo. Honnell, Wolffenstein........143–8; disc. 148
New Haven Speed-Merchant Locomotive, Design and Analysis of the Motor-Alternator Speed Regulator for the. McElhenny........10–16
New York, New Haven and Hartford Railroad, Multiple-Unit-Rectifier Motive Power—Inductive Coordination Considerations on the. Hibbard, Garry, Loomis........416–25; disc. 425
Noise in Motor-Driven Home Appliances, Some Problems and Design Factors Concerning. Cunningham, Wolfert........17–22
Noninteracting Controls in Linear Multivariable Systems. Kavanagh........95–9; disc. 99
Nonlinear Feedback Control Systems, An Analytic Method for Finding the Closed Loop Frequency Response of. Ogata........277–85
Nonlinearities in 2-Stage Electrohydraulic Control Valves, Generalized Charts of the Effects in. Zaborszky, Harrington........401–08; disc. 408
Nonlinearities Present in 2-Stage Electrohydraulic Control Valves, A Describing Function for the Multiple. Zaborszky, Harrington........394–401; disc. 408

O

Oil Refinery Distribution System, A Transient Stability and Voltage Study for a Modern. Boice, Durand, Dalasta........159–64
Oil-Well Drilling, Application of D-C Machines to. Hefner........345–50; disc. 350
Optimization of Relay Servos by Use of Discontinuous Damping, Quasi-. Harris, McDonald, Thaler........292–6; disc. 296

P

Performance Criteria, The Selection and Use of Servo. Schultz, Rideout........383–7; disc. 387
Permanent Magnet Alternators for High Speed Drive, Salient-Pole. Puder, Strauss........333–8
(Petroleum Industry) Application of D-C Machines to Oil-Well Drilling. Hefner........345–50; disc. 350
(Petroleum Industry) Digital Telemetering and Print Out of Pipe Line Information. Nigh........119–22; disc. 122
Pilot Control-Stick Motion, Informative Feedback in Jet-. Diamantides........243–9
Pilotless Aircraft, Silver-Zinc Batteries as Source Primary Electric Power for. Abrahamson....297–300
Pipe Line Information, Digital Telemetering and Print Out of. Nigh........119–22; disc. 122
Poles of Control Systems, A Quick Method for Estimating Closed-Loop. Chen........80–6; disc. 86
Pot-Lines, Direct Current Metering of Large Electrochemical. Reagan........300–02
Power Equipment for the Support of Air Force Weapons, Ground-. Holliday........237–43
Power Plants, A 10-Kw Germanium Rectifier for Automatic. Hake........361–6
Power Rectifier Equipments, Silicon. Marshall, Stewart........366–9
Power Supply for Auxiliaries, The Lightweight Trains—Its. Swarner........409–15; disc. 415
Power System Reliability, Economic Evaluation of Industrial. Dickinson........264–9; disc. 269
Power Systems, A Method of Fault Analysis for A-C Aircraft Electric. Rindt, Jessee........338–44
Power Systems and Resulting Lamp Flicker, Survey of Arc-Furnace Installations on. (Committee Report)........170–81; disc. 181
Power Systems, Static Control and Protection of Aircraft Electric. Hulsey, Schuh........209–13
Power Systems Supplying Arc-Furnace Loads, Selection of Buffer Reactors and Synchronous Condensers on. Concordia, Levoy, Thomas.....123–34; disc. 134
Power Systems, Voltage Modulation on Aircraft. Klokow, Yohe........314–19
Power Systems with Inverter Loads, The Stability of Aircraft D-C. Thomas........183–8
Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines. Robinson........192–9
Print Out of Pipe Line Information, Digital Telemetering and. Nigh........119–22; disc. 122
Problems and Design Factors Concerning Noise in Motor-Driven Home Appliances, Some. Cunningham, Wolfert........17–22
Propeller Control Design, The Role of Analog Computers in. Frick........257–63; disc. 263
Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips. Brightman, McGee, Reifschneider........114–19
Protection of Aircraft Electric Power Systems, Static Control and. Hulsey, Schuh........209–13
Protective Systems, Some Applications of Magnetic Amplifiers in Aircraft Generator. Plette, Butler........427–33
Pump on Electric Utility Systems, Load and Economic Aspects of the Residential Heat. Bary........49–53; disc. 53

Q

Quasi-Optimization of Relay Servos by Use of Discontinuous Damping. Harris, McDonald, Thaler........292–6; disc. 296

R

Railroad Electrification for the American Railroads, The Case for a Standard System of. Perkinson........101–04; disc. 104
Railroad, Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven, and Hartford. Hibbard, Garry, Loomis........416–25; disc. 425
Railway, The Electric System of the Rectifier-Type Locomotives for the Virginian. Brown.....68–73
Ram-Air Cooling Systems for Aircraft Generators. Moroney........217–21
Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads, Selection of Buffer. Concordia, Levoy, Thomas........123–34; disc. 134
Rectifier, A High-Current-Density Selenium. Geib, Brown........358–60
Rectifier Applications in Automotive Vehicles, Selenium. Szabo........369–71
Rectifier Equipment for Electrolytic Processes, Germanium. Miller, Hodgson........353–7
Rectifier Equipments, Silicon Power. Marshall, Stewart........366–9
Rectifier for Automatic Power Plants, A 10-Kw Germanium. Hake........361–6
Rectifier Locomotives, Design of Ignitron Tubes for. Belnap, Zehner, Zuvers........214–17
Rectifier Motive Power—Inductive Coordination Considerations on the New York, New Haven and Hartford Railroad, Multiple-Unit-. Hibbard, Garry, Loomis........416–25; disc. 425
Rectifier to a Cyclotron, Application of a Mechanical. Warren........273–6; disc. 277
Rectifier-Type Locomotives for the Viginian Railway, The Electric System of the. Brown........68–73
Reduction of Air-Borne Transformers, Size. Lee.372–6
Refinery Distribution System, A Transient Stability and Voltage Study for a Modern Oil. Boice, Durand, Dalasta........159–64
Regulator for the New Haven Speed-Merchant Locomotive, Design and Analysis of the Motor-Alternator Speed. McElhenny........10–16
Relay Amplifiers with Feedback, The Response of. Gibson, Tuteur........303–07
Relay Servos by Use of Discontinuous Damping, Quasi-Optimization of. Harris, McDonald, Thaler........292–6; disc. 296
Reliability, Economic Evaluation of Industrial Power-System. Dickinson........264–9; disc. 269
Residence for Data Purposes, Heat Pump and Heating Cables Installed in the Same. Jones, Linden........307–14
Residential Air Conditioning, Single-Phase Versus 3-Phase Service for. Anderson, Hutchinson........61–4; disc. 64
Residential Heat Pump on Electric Utility Systems, Load and Economic Aspects of the. Bary........49–53; disc. 53
Response of Relay Amplifiers with Feedback, The. Gibson, Tuteur........303–07
Role of Analog Computers in Propeller Control Design, The. Frick........257–63; disc. 263
Rolling Mills, A Digital Flying Extensometer for Temper. Wells........376–8

S

Salient-Pole Permanent Magnet Alternators for High Speed Drive. Puder, Strauss........333–8
Salient Poles in a Blast-Cooled Alternator, Surface Heat-Transfer Coefficients of. Robinson, Tse...199–204
Sampled System by a Multirate Controller, Compensation of an Error-. Kranc........149–58; disc. 158
Saturated Servomechanisms, Extended Switching Criterion for Second-Order. Diesel........388–93
Second-Order Saturated Servomechanisms, Extended Switching Criterion for. Diesel........388–93
Selection and Use of Servo Performance Criteria, The. Schultz, Rideout........383–7; disc. 387
Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads. Concordia, Levoy, Thomas........123–34; disc. 134
Selenium Rectifier, A High-Current-Density. Geib, Brown........358–60
Selenium Rectifier Applications in Automotive Vehicles. Szabo........369–71
Series Trips, Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker. Brightman, McGee, Reifschneider........114–19

Servo Network, Transfer Matrix Stability Criterion Applied to an Amplidyne. Honnell, Wolffenstein. ....143–8; disc. 148
Servo Performance Criteria, The Selection and Use of. Schultz, Rideout. ....383–7; disc. 387
Servomechanisms, A New Method for Compensating Contactor. Embler, Weaver....165–70
Servomechanisms, Adaptive. Drenick, Shahbender....286–91; disc. 291
Servomechanisms by a Graphical Method, A New Stability Criterion of Linear. Numakura, Miura....40–8
Servomechanisms, Extended Switching Criterion for Second-Order Saturated. Diesel....388–93
Servos by Use of Discontinuous Damping, Quasi-Optimization of Relay. Harris, McDonald, Thaler....292–6; disc. 296
Shapes Found in Industrial Power Engineering, Analysis of Irregular Current and Voltage Wave. Kaufmann....1–5
Signals, Evaluation of Feedback Control Systems Subjected to Large. Haas....88–95
Silicon Power Rectifier Equipments. Marshall, Stewart....366–9
Silver-Zinc Batteries as Source Primary Electric Power for Pilotless Aircraft. Abrahamson....297–300
Single-Phase Transmission Lines, 400 to 20,000 Cycles, Criteria for Industrial Application of. Munson, Germain....379–82
Single-Phase Versus 3-Phase Service for Residential Air Conditioning. Anderson, Hutchinson....61–4; disc. 64
Size Reduction of Air-Borne Transformers. Lee..372–6
Some Applications of Magnetic Amplifiers in Aircraft Generator Protective Systems. Plette, Butler....427–33
Source-Primary Electric-Power for Pilotless Aircraft, Silver-Zinc Batteries as. Abrahamson....297–300
Speed Regulator for the New Haven Speed-Merchant Locomotive, Design and Analysis of the Motor-Alternator. McElhenny....10–16
Stability Criterion of Linear Servomechanisms by a Graphical Method, A. Numakura, Miura...40–8
Stability of Aircraft D-C Power Systems with Inverter Loads, The. Thomas....183–8
Standard System of Railroad Electrification for the American Railroads, The Case for a. Perkinson....101–04; disc. 104
Static Control and Protection of Aircraft Electric Power Systems. Hulsey, Schuh....209–13
Steel Mill Control Systems, Modern Computer Analysis for the Design of. Reider, Spergel....105–09; disc. 109
Study for a Modern Oil Refinery Distribution System, A Transient Stability and Voltage. Boice, Durand, Dalasta....159–64
Study of the Transfer Function of Contact-Modulated Amplifiers, A. Krantz, Salati, Berkowitz....23–8
Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator. Robinson, Tse..199–204
Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker. (Committee Report)....170–81; disc. 181
Switching Criterion for Second-Order Saturated Servomechanisms, Extended. Diesel....388–93
Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads, Selection of Buffer Reactors and. Concordia, Levoy, Thomas....123–34; disc. 134
Synthesis Method for Multipole Control Systems, A. Freeman....28–31
System by a Multirate Controller, Compensation of an Error-Sampled. Kranc....149–58; disc. 158
Systems, Development of Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric. Markowitz....222–32
Systems for Aircraft Generators, Ram-Air Cooling. Moroney....217–21
Systems, Noninteracting Controls in Linear Multivariable. Kavanagh....95–9; disc. 99
Systems Subjected to Large Signals, Evaluation of Feedback-Control. Haas....88–95

T

Telemetering and Print Out of Pipe-Line Information, Digital Nigh....119–22; disc. 122
Temper Rolling Mills, A Digital Flying Extensometer for. Wells....376–8
Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines, Prediction of Transient and Equilibrium Component. Robinson....192–9
10-Kw Germanium Rectifier for Automatic Power Plants, A. Hake....361–6
Thermal Consideration of Generators in High-Speed Aircraft. Kusko....205–08
380- to 420-Cps Aircraft Electric Systems, Development of Definitions and Limits for Frequency Transients and Frequency Modulation in. Markowitz.222–32
Three-Phase Service for Residential Air Conditioning, Single-Phase Versus. Anderson, Hutchinson....61–4; disc. 64
Train—Its Power Supply for Auxiliaries, The Lightweight. Swarner....409–15; disc. 415
Trains, Diesel Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger. Steinbrook....136–42
Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator, Surface Heat-. Robinson, Tse.199–204
Transfer Function of Contact-Modulated Amplifiers, A Study of the. Krantz, Salati, Berkowitz....23–8
Transfer Matrix Stability Criterion Applied to an Amplidyne Servo Network. Honnell, Wolffenstein....143–8; disc. 148
Transformers, Size Reduction of Air-Borne. Lee....372–6
Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines, Prediction of. Robinson....192–9
Transient Stability and Voltage Study for a Modern Oil Refinery Distribution System, A. Boice, Durand, Dalasta....159–64
Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems, Development of Definitions and Limits for Frequency. Markowitz....222–32
Transmission for Lightweight Passenger Trains, Diesel Locomotives with Mechydro Hydraulic. Steinbrook....136–42
Transmission Lines, 400 to 20,000 Cycles, Criteria for Industrial Application of Single-Phase. Munson, Germain....379–82
(Transportation, Land) An 8,500-Hp Gas-Turbine-Electric Locomotive. Smith....5–10
(Transportation, Land) Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive. McElhenny....10–16
(Transportation, Land) Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad. Hibbard, Garry, Loomis....416–25; disc. 425
(Transportation, Land) The Electric System of the Rectifier-Type Locomotives for the Virginian Railway, Brown....68–73
(Transportation, Land) The Leightweight Train—Its Power Supply for Auxiliaries. Swarner....409–15; disc. 415
Tubes for Rectifier Locomotives, Design of Ignitron. Belnap, Zehner, Zuvers....214–17
Turboprop Aircraft, Design of Generating Equipment for Electra. Alger, Warden, Hansen, Klein.320–32
20,000 Cycles, Criteria for Industrial Application of Single-Phase Transmission Lines, 400 to. Munson, Germain....379–82
Two-Dimensional Feedback Control System, A. Sarachik, Ragazzini....55–60; disc. 60

U

Utility Systems, Load and Economic Aspects of the Residential Heat Pump on Electric. Bary....49–53; disc. 53

V

Valves, A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control. Zaborszky, Harrington....394–401; disc. 408
Valves, Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control. Zaborszky, Harrington....401–08; disc. 408
Virginian Railway, The Electric System of the Rectifier-Type Locomotives for the. Brown....68–73
Voltage Modulation on Aircraft Power Systems. Klokow, Yohe....314–19
Voltage Study for a Modern Oil Refinery Distribution System, A Transient Stability and. Boice, Durand, Dalasta....159–64

W

Wave Shapes Found in Industrial Power Engineering; Analysis of Irregular Current and Voltage. Kaufmann....1–5
Weapons, Ground-Power Equipment for the Support of Air Force. Holliday....237–43
Work-Coil Design, Flat. Mathews....249–56

Z

Zinc Batteries as Source Primary Electric Power for Pilotless Aircraft, Silver-. Abrahamson...297–300

2. Author Index

Abrahamson, Lyal T. Silver-Zinc Batteries As Source-Primary Electric Power for Pilotless Aircraft. (57–491)....297–300
Alger, J. R. M.; W. E. Warden, W. O. Hansen, L. Klein. Design of Generating Equipment for Electra Turboprop Aircraft. (57–471)....320–32
Ambrosius, Carl C.; R. H. Sarikas. Disc....65
Anderson, A. S.; C. Hutchinson. Single-Phase Versus 3-Phase Service for Residential Air Conditioning. (57–6)....61–4; disc. 67
Axelrod, L. R.; W. L. Kinney. Disc....408

B

Backenstoss, H. B. Disc....270
Baker, R. M. Design and Calculation of Induction-Heating Coils. (57–4)....31–8; disc. 39
Barnes, E. C. Disc....104
Bary, Constantine W. Load and Economic Aspects of the Residential Heat Pump on Electric Utility Systems. (57–12)....49–53; disc. 54
Bassett, H. L. Disc....80
Belnap, B. E.; J. L. Zehner, H. E. Zuvers. Design of Ignitron Tubes for Rectifier Locomotives. (57–33)....214–17
Berkowitz, R. S.; F. H. Krantz, O. M. Salati. A Study of the Transfer Function of Contact-Modulated Amplifiers. (57–204)....23–8
Bertram, J. E. Disc....158
Biernson, George. Disc....86
Bills, G. W. Disc....86
Binney, E. A. Disc....425
Biringer, Paul P. Disc....39
Boice, C. W.; S. R. Durand, D. Dalasta. A Transient Stability and Voltage Study for a Modern Oil Refinery Distribution System (56–767)....159–64
Boice, W. K. Disc....134
Brady, Bryce; B. Mohr. Disc....67
Brereton, D. S. Disc....270
Brightman, F. P.; R. R. McGee, P. F. Reifschneider. Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips. (56–652)....114–19
Brown, H. F. Disc....415, 426
Brown, J. C. The Electric System of the Rectifier Type Locomotives for the Virginian Railway. (57–32)....68–73
Brown, W. E.; C. C. Geib. A High-Current-Density Selenium Rectifier. (57–1174)....358–60
Browning, E. H. Disc....110
Butler, J. W.; D. L. Plette. Some Applications of Magnetic Amplifiers in Aircraft Generator Protective Systems. (57–481)....427–33

C

Campbell, H. E. Disc....66
Chen, Kan. A Quick Method for Estimating Closed-Loop Poles of Control Systems. (57–182)....80–6; disc. 87
Concordia, C.; L. G. Levoy. Disc....79, 181
Concordia, C.; L. G. Levoy, C. H. Thomas. Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads. (57–56)....123–34; disc. 134
Crawford, B. E. Disc.... 269
Cunningham, E. R.; G. L. Wolfert. Some Problems and Design Factors Concerning Noise in Motor-Driven Home Appliances. (57–27)....17–22
Curtis, L. B. Disc....104

D

Dalasta, D.; C. W. Boice, S. R. Durand. A Transient Stability and Voltage Study for a Modern Oil Refinery Distribution System. (56–767)....159–64
Derr, W. A. Disc....122
Diamantides, N. D. Informative Feedback in Jet-Pilot Control-Stick Motion. (57–494)....243–9
Dickinson, W. H. Economic Evaluation of Industrial Power System Reliability. (57–639) 264–9; disc. 269
Diehl, A. E. Disc....65
Diesel, J. W. Extended Switching Criterion for Second-Order Saturated Servomechanisms....388–93
Doda, Casimir J. Disc....86
Dortort, I. K. Disc....277
Drenick, R. F.; R. A. Shahbender. Adaptive Servomechanisms. (57–388)....286–91; disc. 291
Durand, S. R.; C. W. Boice, D. Dalasta. A Transient Stability and Voltage Study for a Modern Oil Refinery Distribution System. (56–767)....159–64

E

Embler, J. N.; C. H. Weaver. A New Method for Compensating Contactor Servomechanisms. (57–393)....165–70
Erlandson, N. H. Disc....67

Vol.
Pt. 2
1957

F

Freeman, H. Disc. 99
Freeman, Herbert. A Synthesis Method for Multipole Control Systems. (57-44) 28-31
Frick, Roy K. The Role of Analog Computers in Propeller Control Design. (57-501) 257-63

G

Gallaher, B. M. Disc. 66
Garry, F. T.; L. J. Hibbard, G. N. Loomis. Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad. (55-207) .... 416-25; disc. 426
Geib, C. C.; W. E. Brown. A High-Current-Density Selenium Rectifier. (57-1174) 358-60
Germain, A. J.; W. A. Munson. Criteria for Industrial Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles. (57-1021) 379-82
Gibson, John E.; F. B. Tuteur. The Response of Relay Amplifiers with Feedback. (57-183) 303-07

H

Haas, Vinton, B., Jr. Evaluation of Feedback-Control Systems Subjected to Large Signals. (57-201) 88-95
Hake, E. A. A 10-Kw Germanium Rectifier for Automatic Power Plants. (57-1175) 361-6
Hansen, W. O.; J. R. M. Alger, W. E. Warden, L. Klein. Design of Generating Equipment for Electra Turboprop Aircraft. (57-471) 320-32
Harper, H. W.; T. R. Macon, A. F. Sedgwick. Arc-Furnace Corrective Equipment Using a High Value of Buffer Reactance. (57-55) .... 74-9; disc. 80
Harrington, H. J.; J. Zaborszky. A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves. (57-779) .... 394-401; disc. 408
Harrington, H. J.; J. Zaborszky. Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control Valves. (57-780) .... 401-08; disc. 409
Harris, W. L., Jr.; C. McDonald, G. J. Thaler. Quasi-Optimization of Relay Servos by Use of Discontinuous Damping. (57-605) 292-6
Hefner, B. H. Application of D-C Machines to Oil-Well Drilling. (57-978) .... 345-50; disc. 352
Hibbard, L. J.; F. T. Garry, G. N. Loomis. Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad. (55-207) .. 416-25; disc. 426
Higgins, S. P. Disc. 387
Hodgson, W. R.; L. G. Miller. Germanium Rectifier Equipment for Electrolytic Processes. (57-1178) 353-7
Hogwood, E. E. Disc. 352
Holliday, T. B. Ground-Power Equipment for the Support of Air Force Weapons. (57-500) .. 237-43
Honnell, Pierre M.; D. Wolffenstein. Transfer Matrix Stability Criterion Applied to an Amplidyne Servo Network. (57-186) .... 143-8; disc. 148
Hulsey, R. E.; N. F. Schuh. Static Control and Protection of Aircraft Electric Power Systems. (57-469) 209-13
Hutchinson, Chase; A. S. Anderson. Single-Phase Versus 3-Phase Service for Residential Air Conditioning. (57-6) .... 61-4; disc. 67

J

Jessee, R. D.; L. J. Rindt. A Method of Fault Analysis for A-C Aircraft Electric Power Systems. (57-479) 338-44
Johnson, L. L. Disc. 351
Johnson, Wayne H. Disc. 65
Jones, Benson M. Disc. 134
Jones, C. W.; E. E. Linden. Heat Pump and Heating Cables Installed in the Same Residence for Data Purposes. (57-539) 307-14

K

Kalman, R. E. Disc. 60
Kaufmann, R. H. Analysis of Irregular Current and Voltage Wave Shapes Found in Industrial Power Engineering. (57-1) 1-5
Kavanagh, R. J. Noninteracting Controls in Linear Multivariable Systems. (57-41) .... 95-9; disc. 100
Kidder, A. H. Disc. 67
Kinney, W. L.; L. R. Axelrod. Disc. 408
Kirk, Joe S. Disc. 113
Klein, L.; J. R. M. Alger, W. E. Warden, W. O. Hansen. Design of Generating Equipment for Electra Turboprop Aircraft. (57-471) 320-32
Klokow, R. E.; C. F. Yohe. Voltage Modulation on Aircraft Power Systems. (57-470) 314-19
Koopman, R. J. W. Disc. 148
Kranc, George M. Compensation of an Error-Sampled System by a Multirate Controller. (57-187) .... 149-58; disc. 155
Krantz, F. H.; O. M. Salati, R. S. Berkowitz. A Study of the Transfer Function of Contact-Modulated Amplifiers. (57-204) 23-8
Kusko, Alexander. Thermal Consideration of Generators in High-Speed Aircraft. (57-463) 205-08

L

Lawrence, Robert F. Disc. 64
Lee, Raye. Size Reduction of Air-Borne Transformers. (57-1013) 372-6
Levoy, L. G.; C. Concordia. Disc. 79, 181
Levoy, L. G.; C. Concordia, C. H. Thomas. Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads. (57-56) .... 123-34; disc. 134
Linden, E. E.; C. W. Jones. Heat Pump and Heating Cables Installed in the Same Residence for Data Purposes. (57-539) 307-14
Lloyd, B. L. Disc. 79, 182
Loomis, G. N.; L. J. Hibbard, F. T. Garry. Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven and Hartford Railroad. (55-207) .... 416-25; disc. 426
Lowe, M. H. Disc. 191

M

Macon, T. R.; H. W. Harper, A. F. Sedgwick. Arc-Furnace Corrective Equipment Using a High Value of Buffer Reactance. (57-55) . 74-9; disc. 80
Markowitz, Oscar. Development of Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems. (57-482) 222-32
Marshall, J. A.; V. N. Stewart. Silicon Power Rectifier Equipments. (57-1176) 366-9
Mathews, B. E. Flat Work-Coil Design. (57-493) 249-56
McCausland, Ian. Disc. 291
McClure, H. L. Disc. 65
McDonald, C.; W. L. Harris, Jr.; G. J. Thaler. Quasi-Optimization of Relay Servos by Use of Discontinuous Damping. (57-605) 292-6
McElhenny, S. W. Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive. (57-15) 10-16
McGee, R. R.; F. P. Brightman, P. J. Reifschneider. Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips. (56-652) 114-19
McGrath, William A. Disc. 65
Miller, L. G.; W. R. Hodgson. Germanium Rectifier Equipment for Electrolytic Processes. (57-1178) 353-7
Miura, T.; T. Numakura. A New Stability Criterion of Linear Servomechanisms by a Graphical Method. (57-185) 40-8
Mohr, Bela; B. Brady. Disc. 67
Montross, R.; D. L. Pettit. An Ignitron Contactor for Direct-Press Drives. (57-191) 110-13
Moroney, Richard M. Ram-Air Cooling Systems for Aircraft Generators. (57-472) 217-21
Munson, W. A. Disc. 134
Munson, W. A.; A. J. Germain. Criteria for Industrial Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles. (57-1021) 379-82

N

Nesline, Frederick W., Jr. Disc. 296
Nigh, Max T. Digital Telemetering and Print Out of Pipe Line Information. (57-38) .... 119-22; disc. 122
Numakura, Toshio; T. Miura. A New Stability Criterion of Linear Servomechanisms by a Graphical Method. (57-185) 40-8

O

Ogata, Katsuhiko. An Analytic Method for Finding the Closed-Loop Frequency Response of Nonlinear Feedback-Control Systems. (57-634) 277-85
Ogden, H. S. Disc. 425
O'Neill, P. F. Disc. 53

P

Peltier, Irving A. Disc. 263
Perkinson, T. F. The Case for a Standard System of Railroad Electrification for the American Railroads. (57-16) 101-04
Pettit, D. L.; R. Montross. An Ignitron Contactor for Direct-Press Drives. (57-191) 110-13
Phillips, R. A. Disc. 109
Plette, D. L.; J. W. Butler. Some Applications of Magnetic Amplifiers in Aircraft Generator Protective Systems. (57-481) 427-33
Puder, Allen T.; F. Strauss. Salient-Pole Permanent Magnet Alternators for High Speed Drive. (57-489) 333-8

Q

Quick, C. E. Disc. 182

R

Ragazzini, J. R.; P. Sarachik. A 2-Dimensional Feedback Control System. (57-184) ... 55-60; disc. 61
Reagan, M. E. Direct Current Metering of Large Electrochemical Pot-Lines. (57-228) 300-02
Reider, J. E.; P. Spergel. Modern Computer Analysis for the Design of Steel Mill Control Systems. (56-875) .... 105-09; disc. 110
Reifschneider, P. J.; F. P. Brightman, R. R. McGee. Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips. (56-652) 114-19
Rideout, V. C.; W. C. Schultz. The Selection and Use of Servo Performance Criteria. (57-1033) .... 383-7; disc. 387
Ries, E. P. Disc. 344
Rindt, L. J.; R. D. Jessee. A Method of Fault Analysis for A-C Aircraft Electric Power Systems. (57-479) 338-44
Robinson, Walter. Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines. (57-460) 192-9
Robinson, W.; F. S. Tse. Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator. (57-461) 199-204

S

Salati, O. M.; F. H. Krantz, R. S. Berkowitz. A Study of the Transfer Function of Contact-Modulated Amplifiers. (57-204) 23-8
Sarachik, Philip; J. R. Ragazzini. A 2-Dimensional Feedback Control System. (57-184) .... 55-60; disc. 61
Sarikas, Robert H.; C. C. Ambrosius. Disc. 65
Schuh, N. F.; R. E. Hulsey. Static Control and Protection of Aircraft Electric Power Systems. (57-469) 209-13
Schultz, W. C.; V. C. Rideout. The Selection and Use of Servo Performance Criteria. (57-1033) .... 383-7; disc. 387
Schwabe, W. E. Disc. 182
Sedgwick, A. F.; H. W. Harper, T. R. Macon. Arc-Furnace Corrective Equipment Using a High Value of Buffer Reactance. (57-55) .... 74-9; disc. 80
Shahbender, R. A.; R. F. Drenick. Adaptive Servomechanisms. (57-388) .... 286-91; disc. 291
Sheheen, S. J. Electric Ignition for Conductive Fuels. (57-492) 233-6
Sklansky, Jack. Disc. 60
Smith, R. E. A Brushless Air-Cooled Aircraft A-C Generator. (57-468) .... 189-91; disc. 191
Smith, R. M. An 8,500-Hp Gas-Turbine-Electric Locomotive. (57-28) 5-10
Spergel, Philip, J. E. Reider. Modern Computer Analysis for the Design of Steel Mill Control Systems. (56-875) .... 105-09; disc. 110
Spooner, Morton G. Disc. 387
Steinbrook, S. E. Diesel Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger Trains. (57-30) 136-42
Stewart, V. N.; J. A. Marshall. Silicon Power Rectifier Equipments. (57-1176) 366-9
Strauss, Fritz; A. T. Puder. Salient-Pole Permanent Magnet Alternators for High Speed Drive. (57-489) 333-8
Swarner, J. L. The Lightweight Train—Its Power Supply for Auxiliaries. (56-256) . 409-51; disc. 415
Szabo, Emery J. Selenium Rectifier Applications in Automotive Vehicles. (57-1177) 369-71

T

Thaler, G. J.; W. L. Harris, Jr., C. McDonald. Quasi-Optimization of Relay Servos by Use of Discontinuous Damping. (57-605) 292-6
Thomas, C. H.; C. Concordia, L. G. Levoy. Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads. (57-56) .... 123-34; disc. 134
Thomas, R. E. The Stability of Aircraft D-C Power Systems with Inverter Loads. (57-478) 183-8

Tse, F. S.; W. Robinson. Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator. (57–461)......................199–204
Tudbury, C. A. Disc..............................39
Tuteur, Frana B.; J. E. Gibson. The Response of Relay Amplifiers with Feedback. (57–183) 303–07

W

Warden, W. E.; J. R. M. Alger, W. O. Hansen, L. Klein. Design of Generating Equipment for Electra Turboprop Aircraft. (57–471).....320–32
Warren, Jack. Application of a Mechanical Rectifier to a Cyclotron. (57–700)..................273–6
Weaver, C. H.; J. N. Embler. A New Method for Compensating Contactor Servomechanisms. (57–393)..................................165–70
Webb, G. W. Disc..............................350
Weed, H. R. Disc..............................39
Wells, N. S. A Digital Flying Extensometer for Temper Rolling Mills. (57–703)..................376–8
Wolfert, G. L.; E. R. Cunningham. Some Problems and Design Factors Concerning Noise in Motor-Driven Home Appliances. (57–27)........17–22
Wolffenstein, Daniel, P. M. Honnell. Transfer Matrix Stability Criterion Applied to an Amplidyne Servo Network. (57–186)..............143–8; disc. 148

Y

Yohe, C. F.; R. E. Klokow. Voltage Modulation on Aircraft Power Systems. (57–470)........314–19

Z

Zaborszky, J.; H. J. Harrington. A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves. (57–779). ....394–401; disc. 408
Zaborszky, J.; H. J. Harrington. Generalized Charts of the Effects in Nonlinearities in 2-Stage Electrohydraulic Control Valves. (57–780)............ ....401–08; disc. 409
Zehner, J. L.; B. E. Belnap, H. E. Zuvers. Design of Ignitron Tubes for Rectifier Locomotives. (57–33) 214–17
Zuvers, H. E.; B. E. Belnap, J. L. Zehner. Design of Ignitron Tubes for Rectifier Locomotives. (57–33) 214–17

Analysis of Irregular Current and Voltage Wave Shapes Found in Industrial Power Engineering

R. H. KAUFMANN
FELLOW AIEE

THE industrial power application engineer is faced more and more with the problem of resolving the properties of irregular wave shapes of current and voltage. This is due to the ever-increasing use of current-limiting fuses, electronic switching devices, magnetic amplifiers, gaseous discharge tubes, lamps, etc. For the most part, the quantities to be resolved are: the current-time product ($\int idt$), and the summation of the current squared-time product ($\int i^2dt$). From the former are derived average, and from the latter root-mean-square values of current. Which of these two is desired in a particular problem will depend upon whether the effects being studied are responsive to the first power or to the second power of the current.

In general, the engineer will be confronted with an irregular-shaped wave, consisting of a single excursion. In many cases, this identical wave shape will be recurrent at regular intervals. The solution nonetheless, requires evaluation of the properties of a single current excursion. The peculiar wave shapes result from abrupt transitions in the electric circuit, caused by ignition of a thyratron by grid control; breakdown of the gas within a tube; or the abrupt melting of a fuse link in a current-limiting fuse. The wave shapes seldom can be expressed as a simple mathematical function of time.

There are a number of well-recognized procedures by which these basic electrical properties can be evaluated. The method proposed in this paper is believed to permit quick evaluation with good practical accuracy with a number of important advantages: 1. It allows measurements to be made directly on the record print as obtained from an oscillograph or a direct transfer tracing paper. No replotting of the wave is required. 2. The results are computed in terms of crest current, total interval time, and a multiplying factor. This makes it easy to visualize mentally the approximate order of the answer and, thus, avoid errors. 3. The procedure is simple and saves time.

Solution by Parts

The unique approach to the problem proposed in this paper is based on subdividing the over-all wave shape into a number of individual sections, each of which will be evaluated separately. The current-time product of the complete wave is simply the sum of the current-time products for each individual section. The same procedure applies to the summation of current-squared-time products. The key to the solution is that of subdividing the wave into individual sections whose geometry corresponds with one whose characteristics are known.

Throughout the paper, reference will be made to waves of current with respect to time. Identically the same procedures are applicable to waves of voltage with respect to time.

Fig. 1 presents a family of geometric configurations of current-time waves with which are associated accurate evaluations of the current-time products, Fig. 1(A) and the current-squared-time products, Fig. 1(B). The family of waves changes progressively from a rectangular shape to a trapezoidal wave shape, to a rounded wave shape produced by a sine wave, to a more peaked rounded shape produced by a one-minus-cosine wave, to a still more peaked figure with straight-line sides forming a triangle, and concluding with a still more peaked shape in which current increases as the square of time. In all cases (except for the trapezoidal figure), the essential qualities are expressed in terms of the crest magnitude of current and the base time duration. The numerical modifier in the case of the summation of current-squared-time progressively diminishes from 1.0 for the rectangular current pattern to 0.2 for the very peaked wave shape.

Derivations of the exact solutions applicable to the unit sections are included in the appendix.

For the case of the trapezoidal wave shape, it will be observed that the average value is easy to establish, whereas the current squared-time product is rather complex. Accordingly, Fig. 2 has been prepared which makes it possible to derive the rms value from the average value by a simple numerical multiplier established by the ratio of the smaller to the higher ordinate of the trapezoid. Thus, the procedure in evaluating the current squared-time product becomes one of first multiplying the average value (easy to compute) by the rms/avg ratio obtained from Fig. 2 which gives the rms current value. This rms current value squared and multiplied by the duration gives the desired current squared-time product.

Procedure

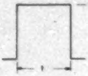
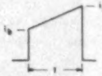
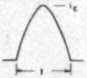
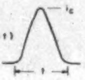
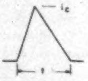
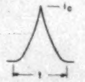
The first step in the procedure is to subdivide the total current or voltage wave shape into a number of sections each of which closely approaches one of the geometric figures whose characteristics are known. Reference to Figs. 3 and 4 will aid in understanding this step. In this example, the actual wave has been subdivided into four sections: A, resembling a sine wave; B, a triangle; C, a rectangle; and D, a triangle. In many practical cases, as illustrated in the example, the individual sections will not match exactly any of the simple geometric patterns. If one is seeking the highest accuracy, it might be desirable further to subdivide the wave into smaller sections. On the other hand, deviations from perfect wave shapes can be allowed and still secure practical answers. Bear in mind always, that it is most important that the figures closely match at high current. If mismatches must exist, allow them to exist at the low-current parts of the figures. This principle is evident in the treatment of the example.

The next step in the procedure involves the computation of electrical properties of each individual section of the over-all wave. In the illustrative example, both the current-time products and the current squared-time products have been computed for completeness. In most specific practical problems, it is likely that only one of these two basic quantities would be needed.

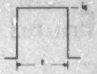
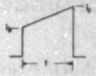
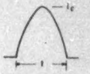
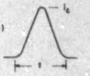
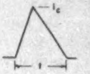
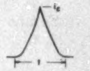
The numerical modifier applicable to the unit sections are obtained from Fig. 1.

Paper 57-1, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted August 20, 1956; made available for printing October 17, 1956.

R. H. KAUFMANN is with the General Electric Company, Schenectady, N. Y.

WAVE SHAPE		THROUGH THE TIME INTERVAL, t	
		I_{AVG}	$\int_0^t i\,dt = (I_{AVG})(t)$
$i = k$		i_c	$i_c t$
$i = kt$		$\frac{i_b + i_c}{2}$	$\frac{1}{2}(i_b + i_c)\,t$
$i = k \sin t$		$\frac{2}{\pi} i_c$	$\frac{2}{\pi} i_c t$
$i = k(1 - \cos t)$		$\frac{1}{2} i_c$	$\frac{1}{2} i_c t$
$i = kt$		$\frac{1}{2} i_c$	$\frac{1}{2} i_c t$
$i = kt^2$		$\frac{1}{3} i_c$	$\frac{1}{3} i_c t$

A—Factors for determining average currents (or voltages)

WAVE SHAPE		THROUGH THE TIME INTERVAL, t	
		I_{RMS}	$\int_0^t i^2 dt = (I_{RMS})^2 (t)$
$i = k$		i_c	$i_c^2 t$
$i = kt$		$\sqrt{\frac{i_b^2 + i_b i_c + i_c^2}{3}}$	$\frac{1}{3}(i_b^2 + i_b i_c + i_c^2)\,t$ *
$i = k \sin t$		$\frac{i_c}{\sqrt{2}}$	$\frac{1}{2} i_c^2 t$
$i = k(1 - \cos t)$		$\frac{1}{2}\sqrt{3/2}\, i_c$	$\frac{3}{8} i_c^2 t$
$i = kt$		$\frac{1}{\sqrt{3}} i_c$	$\frac{1}{3} i_c^2 t$
$i = kt^2$		$\frac{1}{\sqrt{5}} i_c$	$\frac{1}{5} i_c^2 t$

* See Fig. 2.

B—Factors for determining rms currents (or voltages)

Fig. 1. **Properties of various familiar wave shapes**

Part A is evaluated as a sine wave, part B as a triangle, part C as rectangle, and part D as a triangle.

The total ampere-time product for the entire wave shape is obtained by summing up these quantities for the four parts, using the factors from Fig. 1(A), giving a total of 20.58 amp-sec (amperes per second). For some problems (for example, to determine how much metal would be deposited electrolytically by this current wave shape), the ampere-time product might be all that is needed. It might be desired to know the average current for the period of current conduction. This would be determined by dividing the total current-time product by the total elapsed time which, in the example, would be 4,780 amp.

The procedure for rms evaluation is quite similar, except that factors from Fig. 1(B) are used. The total current squared-time product would be resolved by summing up this quantity for the four separate parts which, for the example, is 36.97 (10^4) amp²-sec. For some problems (for instance, to determine the increase in temperature of a conductor due to this current flow), this current-squared-time product alone would suffice. On the other hand, should it be desired to determine the rms current for the time interval of conduction, this would be obtained by dividing the sum total current squared-time product by the total elapsed time and extracting the square root. For the example problem, the answer would be 9,270 amp.

Repeated Waves

The problem of resolving a succession of identical irregular wave shapes of current or voltage separated by equal time intervals involves only one additional step. The properties of one individual wave pattern would be resolved in the manner just described. Associated with such a repeating chain of waves is a steady-state average value of current and an rms value of current. An illustrative pattern of repeating wave shapes appears in Figs. 5 and 6. First, there is a pattern with unit waves repeating with the same polarity at 1/60-sec intervals. The steady-state quantities are resolved using one complete unit cycle time of 1/60 second. Using the data previously evaluated for a unit wave shape, the steady-state rms current would be determined by

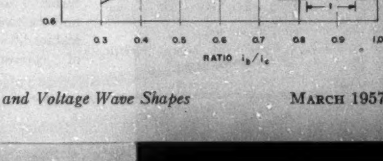


Fig. 2. Factors for determining rms/average ratio of a trapezoidal current (voltage) wave

Through the time interval (t)
rms = i_c (rms factor)
avg = i_c (avg factor)

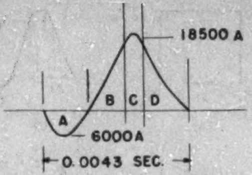
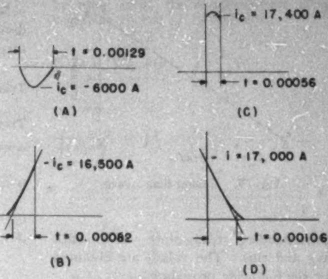


Fig. 3. Complete wave to be analyzed

Fig. 4 (right). Parts of wave shown in Fig. 3

A—Sine wave
B—Triangle
C—Rectangle
D—Triangle



dividing the total current-squared-time product of one wave by the unit cycle time of 0.0167 sec and then extracting the square root giving, for the effective current, an answer of 4,700 amp. The steady-state average current is obtained by dividing the total current-time product of one wave by the unit cycle time of 0.0167 sec, giving an answer of 1,230 amps.

Should the wave shape be repeated with alternately reversed polarity in the manner illustrated in Fig. 6, the evaluation procedure is affected as follows. The rms current values are always positive and uninfluenced by polarity. In contrast, the current-time products of the successive waves would have opposite polarity signs and would add up to zero over any integral number of unit cycles. The corresponding average current value would thus also be zero.

Conclusions

The procedures presented in this paper are accurate to the extent that the sectional parts of the wave shape being analyzed agree with the ideal wave shape. By positioning the ideal geometric pattern so that agreement exists at the high-current part of the figure, good accuracy can be achieved even though there appears to be considerable deviation in the low-current portion. The procedures herein described have been used by the author for several years in evaluating the properties of irregular wave shapes of both current and voltage and have been found quite simple to apply and to be adequately accurate.

Examples

ANALYSIS IN PARTS

The complete wave is shown in Fig. 3.

Part A. Sine wave, Fig 4(A):

$$\int i\,dt = \frac{2}{\pi}\, i_c t = 0.637(-6{,}000)(0.00129) = 4.93 \text{ amp-sec}$$

$$\int i^2 dt = \frac{1}{2}\, i_c^2 t = \frac{(-6{,}000)^2(0.00129)}{2} = 2.32(10^4) \text{ amp}^2\text{-sec}$$

Part B. Triangle, Fig. 4(B):

$$\int i\,dt = \frac{1}{2}\, i_c t = \frac{(16{,}500)(0.00082)}{2} = 6.77 \text{ amp-sec}$$

$$\int i^2 dt = \frac{1}{3}\, i_c^2 t = \frac{(16{,}500)^2(0.00082)}{3} = 7.45(10^4) \text{ amp}^2\text{-sec}$$

Part C. Rectangle Fig 4(C) (Use average ordinate):

$$\int i\,dt = i_c t = 17{,}400(0.00056) = 9.74 \text{ amp-sec}$$

$$\int i^2 dt = i_c^2 t = (17{,}400)^2(0.00056) = 1.7(10^5) \text{ amp}^2\text{-sec}$$

Part D. Triangle, Fig. 4(D):

$$\int i\,dt = \frac{1}{2}\, i_c t = \frac{(17{,}000)(0.00106)}{2} = 9.0 \text{ amp-sec}$$

$$\int i^2 dt = \frac{1}{3}\, i_c^2 t = \frac{(17{,}000)^2(0.00106)}{3} = 1.02(10^5) \text{ amp}^2\text{-sec}$$

AVERAGE EVALUATION

ΣIt

Part A:	4.93 amp-sec
Part B:	6.77
Part C:	9.74
Part D:	9.0
	20.58 amp-sec

In many cases, this will be the result wanted.

The average current for the interval of current flow ($t = 0.0043$) would be

$$I_{avg} \text{ (total time} = 0.0043) = \frac{\Sigma It}{t} = \frac{20.58}{0.0043} = 4{,}780 \text{ amp}$$

$\Sigma I^2 t$

Part A:	2.32 (10^4) amp²-sec
Part B:	7.45 (10^4)
Part C:	17.0 (10^4)
Part D:	10.2 (10^4)
	36.97 (10^4) amp²-sec

In many cases, this result will be all that is needed.

To determine the rms current for the actual interval of current flow (total $t = 0.0043$ sec)

$$I_{rms} \text{ (total } t = 0.0043) = \sqrt{\frac{\Sigma I^2 t}{t}} = \sqrt{\frac{36.97\ (10^4)}{0.0043}} = 9{,}270 \text{ amp}$$

If, however, the rms current based on a time interval of one cycle of a 60-cycle a-c system, the result would be

$$I_{rms} \text{ (total } t = 0.0167) = \sqrt{\frac{\Sigma I^2 t}{t}} = \sqrt{\frac{36.97\ (10^4)}{0.0167}} = 4{,}700 \text{ amp}$$

REPEATING WAVES

If the wave pattern is regularly repeated at uniform time intervals, there will be

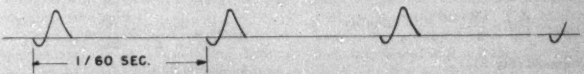


Fig. 5. Repeating wave at uniform time intervals

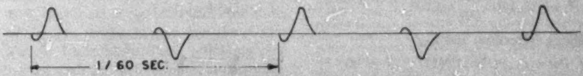


Fig. 6. Repeating wave at half-cycle intervals

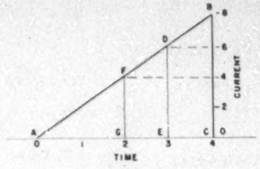


Fig. 7. Current-time wave

corresponding steady-state values, both avg and rms. The values are obtained by the following procedure.

Imagine the wave to be repeated, as shown in Fig. 5.

$$I_{rms}\text{ (steady state)}=\sqrt{\frac{\Sigma I^2t\text{ during one cycle}}{t\text{ one cycle}}}$$

$$=\sqrt{\frac{36.97(10^4)}{0.0167}}$$
$$=4{,}700\text{ amp}$$

$$I_{avg}\text{ (steady state)}=\frac{20.58}{0.0167}=1{,}230\text{ amp}$$

Now suppose the waves to be repeating at half-cycle intervals, as in Fig. 6.

$$I_{rms}\text{ (steady state)}=\sqrt{\frac{\Sigma I^2t\text{ during one cycle}}{t\text{ one cycle}}}$$

$$=\sqrt{\frac{36.97(10^4)+36.97(10^4)}{0.0167}}$$
$$=6{,}650\text{ amp}$$

$$I_{avg}\text{ (steady state)}=\frac{\Sigma It}{t}$$

$$=\frac{20.58-20.58}{0.0167}=0$$

APPROXIMATION

Consider the current-time wave shown in Fig. 7. Assume the problem is to determine I^2t. As the figure is a perfect triangle, it can be exactly evaluated in one part.

$$\int i^2dt=\frac{i_c^2t}{3}=\frac{(8)^2(4)}{3}=85.333\text{ exact}$$

Suppose for analysis, the curve is divided into the triangle, *ADE*, plus trapezoid, *DBCE*.

Triangle $\int i^2dt=\frac{i_c^2t}{3}=\frac{(6)^2(3)}{3}$ $=36.00$

Trapezoid (Fig. 2)

$i_b=6,\ i_c=8,\ i_{avg}=7$

$i_b/i_c=0.75$, ratio $\frac{\text{rms}}{\text{avg}}=1.003$

(ignore correction; use i_{avg})

$\int i^2dt=(i_{avg})^2(t)=(7)^2(1)$ $=49.00$

85.00

Suppose for analysis, the curve is divided into the triangle, *AFG*, plus trapezoid, *FBCG*.

Triangle $\int i^2dt=\frac{i_c^2t}{3}=\frac{(4)^2(2)}{3}=$ 10.67

Trapezoid

$i_b=4,\ i_c=8,\ i_{avg}=6$

$i_b/i_c=0.5$, ratio $\frac{\text{rms}}{\text{avg}}=1.018$

$i_{rms}=i_{avg}(1.018)=6(1.018)=6.108$

$\int i^2dt=(I_{rms})^2(t)=(6.108)^2(2)=$ 74.62

85.29

Appendix I. Computation Fundamentals

Case A, Fig. 8(A). $\mathbf{i=i_c}$

Average:

$$\int idt=\int_0^{t_1} i_cdt=i_c\Big|_0^{t_1}=i_ct_1$$

$$I_{avg}=\frac{\int idt}{t}=\frac{i_ct_1}{t_1}=i_c$$

Root-mean-square:

$$\int i^2dt=\int_0^{t_1}i_c^2dt=i_c^2t\Big|_0^{t_1}=i_c^2t_1$$

$$I_{rms}=\sqrt{\frac{i^2dt}{t}}=\sqrt{\frac{i_c^2t_1}{t_1}}=i_c$$

Case B, Fig. 8(B). $\mathbf{i=i_b+\left(\frac{i_c-i_b}{t_1}\right)t}$

Average:

$$\int idt=\int_0^{t_1}\left[i_b+\left(\frac{i_c-i_c}{t_1}\right)t\right]dt$$

$$=i_bt\Big|_0^{t_1}+\left(\frac{i_c-i_b}{t_1}\right)\frac{t^2}{2}\Big|_0^{t_1}$$

$$=\left(\frac{i_b+i_c}{2}\right)t_1$$

$$I_{avg}=\frac{\int idt}{t}=\left(\frac{i_b+i_c}{2}\right)t_1\div t_1$$

$$=\frac{i_b+i_c}{2}$$

Root-mean-square:

$$\int i^2dt=\int_0^{t_1}\left[i_b+\left(\frac{i_c-i_b}{t_1}\right)t\right]^2dt$$

$$=i_b^2t\Big|_0^{t_1}+\frac{2i_b(i_c-i_b)}{t_1}\,\frac{t^2}{2}\Big|_0^{t_1}+$$
$$\frac{(i_c-i_b)^2}{t_1^2}\,\frac{t^3}{3}\Big|_0^{t_1}$$

$$=(i_b^2+i_bi_c+i_c^2)\,\frac{t_1}{3}$$

$$I_{rms}=\sqrt{\frac{i^2dt}{t}}=\sqrt{\frac{i_b^2+i_bi_c\,i_c^2}{3}}$$

See Fig. 2 for way to derive I_{rms} from I_{avg}:

$$I_{rms}=I_{avg}\times\left[\text{ratio }\frac{\text{rms}}{\text{avg}}\right]$$

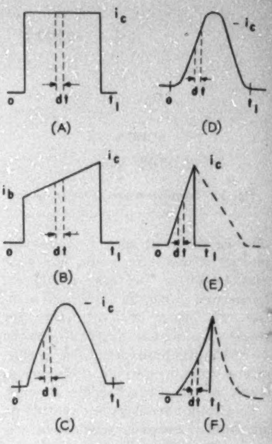


Fig. 8. Computation fundamentals

Case C, Fig. 8(C). $\mathbf{i=i_c\sin\frac{\pi}{t_1}t}$

Average:

$$\int idt=\int_0^{t_1} i_c\sin\frac{\pi}{t_1}tdt$$

$$=\frac{i_ct_1}{\pi}\left(-\cos\frac{\pi}{t_1}t\right)\Big|_0^{t_1}$$

$$=\frac{2}{\pi}i_ct_1$$

$$I_{avg}=\frac{\int idt}{t}=\frac{2}{\pi}i_c$$

Root-mean-square:

$$\int i^2dt=\int_0^{t_1}i_c^2\sin^2\frac{\pi}{t_1}tdt$$

$$=i_c^2\left(\frac{t}{2}\Big|_0^{t_1}-\frac{\sin\frac{2\pi}{t_1}t}{\frac{4\pi}{t_1}}\Bigg|_0^{t_1}\right)$$

$$=\frac{i_c^2t_1}{2}$$

$$I_{rms}=\sqrt{\frac{\int i^2dt}{t}}=\sqrt{\frac{i_c^2t_1}{2t_1}}$$

$$=\frac{i_c}{\sqrt{2}}$$

I_{avg} and I_{rms} will be identical for one-half the wave shown; that is, for a time, $t_1/2$. $\int idt$ and i^2dt would be halved.

Case D, Fig. 8(D). $\mathbf{i=\frac{i_c}{2}\left(1-\cos\frac{2\pi}{t_1}t\right)}$

Average:

$$\int i\,dt = \frac{i_c}{2}\int_0^{t_1}\left(1-\cos\frac{2\pi}{t_1}t\right)dt$$

$$= \frac{i_c}{2}\left[t\Big|_0^{t_1} - \frac{t_1}{2\pi}\left(\sin\frac{2\pi}{t_1}t\right)\Big|_0^{t_1}\right]$$

$$= \frac{i_c t_1}{2}$$

$$I_{avg} = \frac{\int i\,dt}{t} = \frac{i_c}{2}$$

Root-mean-square:

$$\int i^2 dt = \frac{i_c^2}{4}\int_0^{t_1}\left(1-2\cos\frac{2\pi}{t_1}+\cos^2\frac{2\pi}{t_1}t\right)dt$$

$$= \frac{i_c^2}{4}\left(t\Big|_0^{t_1} - \frac{2t_1}{2\pi}\sin\frac{2\pi}{t_1}t\Big|_0^{t_1} + \frac{t}{2}\Big|_0^{t_1} + \frac{\sin\frac{4\pi}{t_1}t}{\frac{8\pi}{t_1}}\Bigg|_0^{t_1}\right)$$

$$= \frac{3}{8}i_c^2 t_1$$

$$I_{rms} = \sqrt{\frac{i^2 dt}{t}} = \sqrt{\frac{3}{8}i_c^2} = \frac{1}{2}\sqrt{\frac{3}{2}}\,i_c$$

Case E, Fig. 8(E). $i = \frac{i_c}{t_1}t$

Average:

$$\int i\,dt = \int_0^{t_1}\frac{i_c}{t_1}t\,dt = \frac{i_c}{t_1}\,\frac{t^2}{2}\Big|_0^{t_1}$$

$$= \frac{i_c t_1}{2}$$

$$I_{avg} = \frac{\int i\,dt}{t} = \frac{i_c t_1}{2} \div t_1 = \frac{i_c}{2}$$

Root-mean-square:

$$\int i^2 dt = \int_0^{t_1}\frac{i_c^2 t^2}{t_1^2}dt$$

$$= \frac{i_c^2}{t_1^2}\left(\frac{t^3}{3}\right)\Big|_0^{t_1} = \frac{i_c^2 t_1}{3}$$

$$I_{rms} = \sqrt{\frac{i^2 dt}{t}} = \sqrt{\frac{i_c^2 t_1}{3} \div t_1} = \frac{1}{\sqrt{3}}\,i_c$$

Since i_c is common to both left and right sections, these computations apply also to the full triangle from zero to crest to zero by using as t_1, the complete time from beginning to end of current flow.

Case F, Fig. 8(F). $i = \frac{i_c}{t_1^2}t^2$

Average:

$$\int i\,dt = \int_0^{t_1}\frac{i_c}{t_1^2}t^2 dt = \frac{i_c}{t_1^2}\left(\frac{t^3}{3}\right)\Big|_0^{t_1}$$

$$= \frac{i_c t_1}{3}$$

$$I_{avg} = \frac{\int i\,dt}{t} = \frac{i_c t_1}{3} \div t_1 = \frac{i_c}{3}$$

Root-mean-square:

$$\int i^2 dt = \int_0^{t_1}\frac{i_c^2}{t_1^4}t^4 dt$$

$$= \frac{i_c^2}{t_1^4}\left(\frac{t^5}{5}\right)\Big|_0^{t_1} = \frac{i_c^2 t_1}{5}$$

$$I_{rms} = \sqrt{\frac{i^2 dt}{t}} = \sqrt{\frac{i_c^2 t_1}{5} \div t_1}$$

$$= \frac{i_c}{\sqrt{5}}$$

An 8,500-Hp Gas-Turbine-Electric Locomotive

R. M. SMITH
ASSOCIATE MEMBER AIEE

AFTER almost a decade of development of the gas-turbine-electric locomotive in practical service, this type of motive power has emerged from the experimental stage. One result of its acceptance in heavy through-freight service has been the demand for a unit considerably more powerful than those previously built. This has been designed and is now under construction. It embodies the same basic design principles found in the smaller units now in service. Wherever possible, advantage has been taken of the lessons already learned to design the new locomotive for increased efficiency, greater reliability, and lower maintenance requirements.

Development

In 1948, the General Electric Company built the first gas-turbine-electric locomotive to be constructed in the United States. This locomotive, shown in Fig. 1, was placed in revenue service of the Union Pacific Railroad where it operated for a period of 21 months, covering a total of approximately 106,000 miles. Its performance during this time was judged highly successful, and an order was placed for ten locomotives (Fig. 2) to be essentially duplicates of the original. In 1952, a second order was placed for 15 locomotives basically similar to the earlier design, but differing considerably in outward appearance, as shown in Fig. 3. Outside walkways were used, rather than cab aisles as on the earlier design. Internally, however, the differences between the two designs were minor.

These 25 locomotives have proved to be particularly adept at handling heavy, high-speed freight service. They are operating between Omaha, Neb., and Ogden, Utah, and account for approximately 10% (per cent) of the total gross ton-miles of freight traffic hauled by the Union Pacific Railroad. Many problems connected with the development and application of such a radically different form of motive power had to be solved. However, the fundamental principle that the gas turbine in single, high-horsepower units can be applied to locomotive service as a prime mover has now been proved by 8 years of successful experience. Accordingly, a new and considerably larger locomotive, shown in Fig. 4, has been designed. It incorporates the lessons gained from the operation of the present gas-turbine locomotive fleet plus the latest advances in the design and construction of gas turbines. Fifteen units are currently under construction.

These new units have a nominal rating of 8,500 hp (horsepower) at 6,000 feet altitude and 90 F (degrees Fahrenheit) ambient. Inasmuch as the performance of the gas turbine is more affected by temperature and altitude than is that of most other prime movers, it is important to note the difference between the rating points of the new and old locomotives fully to appreciate the increase in capacity that is being achieved. The existing locomotives are rated 4,500 hp at 1,500 feet altitude and 80 F. If the same rating

---

Paper 57-28, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21-25, 1957. Manuscript submitted October 22, 1956; made available for printing November 27, 1956.

R. M. SMITH is with the General Electric Company, Erie, Pa.

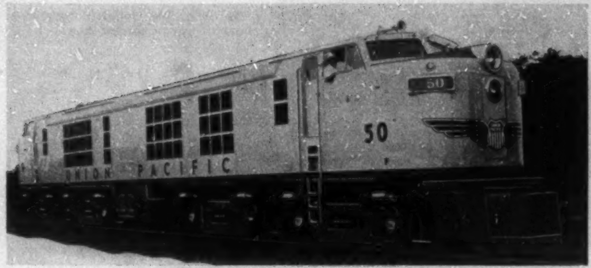


Fig. 1. Pilot gas-turbine-electric locomotive ready for field tests

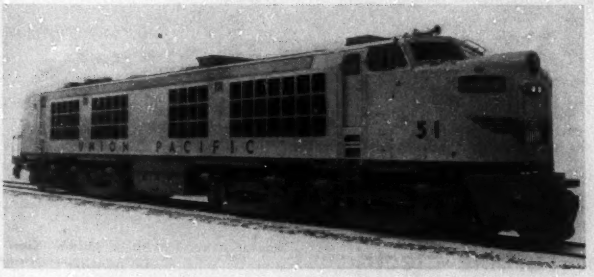


Fig. 2. First design of commercial gas-turbine-electric locomotive

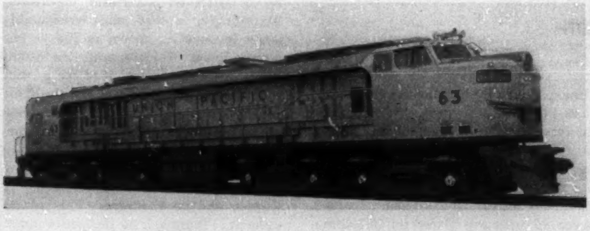


Fig. 3. Second design of commercial gas-turbine-electric locomotive

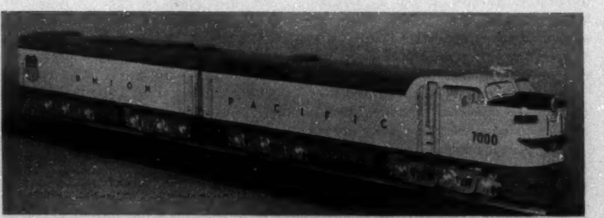


Fig. 4. Model showing general appearance of the 2-unit 8,500-hp locomotive currently under construction

point were used for the 8,500-hp unit, the nominal rating would become 10,700 hp. However, electrical capacity has been installed to handle only 8,500 hp between 18 and 40 mph. Obviously, a generous reserve of power exists in the turbine design, and this should translate itself into lower maintenance costs.

Design Objectives

If a locomotive is to operate consistently as a single unit, a high premium must be placed upon the reliability of each critical device. In many cases it is possible and desirable to provide some of the feature of multiple-unit operation. For example, this has been done in the case of air compressors where two have been installed, either one of which is large enough to handle a normal-size train. Also, the main generators are electrically independent. Generally speaking, however, the opportunities for this approach are limited, and reliability must be gained by reducing the number of components to an absolute minimum and making them of the highest quality possible. Another reason for keeping the number of devices to an absolute minimum is the resulting reduction in the routine maintenance required.

In order to utilize the characteristics of turbine to best advantage, the locomotives must average very high monthly mileages in service. To make this possible, a fuel tender with a capacity of about 23,000 gallons, greater sand capacity, and generous diesel-fuel capacity have been incorporated in the design. In addition, minimizing the number of devices on the locomotive will contribute toward this objective by reducing the amount of out-of-service maintenance time. A comparison of the figures in Table I gives some indication of the progress that has been made in this respect.

Any tabulation of this type can only give an approximate comparison since motors and other items of equipment having vastly different ratings and requirements are necessarily compared in a simple numerical count. It does, however, give an indication of the relative amount of maintenance required by the auxiliary electric equipment.

Gas Turbine

No discussion of the 8,500-hp gas-turbine-electric locomotive could be complete without a description of the turbine itself. The turbine utilizes a simple-cycle arrangement with a 16-stage compressor followed by ten combustion

Table I. Locomotive Items Requiring Maintenance

Item	Five Diesel Units, 8,750 Hp	Previous Gas Turbine, 4,500 Hp	New Gas Turbine, 8,500 Hp
1. Traction motors and generators	25	12	16
2. Auxiliary machines	40	32	15
3. Relays	110	52	24
4. Contactors	70	54	57
5. Reversers and braking switches	10	4	8

chambers and a 2-stage turbine, as shown in Fig. 5. All shaft power is brought out from the compressor end of the turbine, in contrast to the previous designs where the power take-off shaft was brought out through the exhaust stack. The shaft goes directly to the main gear box where the speed is reduced for use by the main traction generators and various turbine auxiliaries.

The turbine is of the so-called 2-bearing design in which the combustion chambers are placed around the outside of the main axial-flow compressor. After leaving the compressor, the air reverses its direction of flow 180 degrees and enters the combustion chambers. Here it again reverses, passing directly from the combustion chambers to the 2-stage turbine, as shown in Fig. 6. This arrangement greatly shortens the over-all turbine length and permits the use of only two main turbine shaft bearings.

The entire turbine casing is split at the shaft center line so that the top half can be removed for servicing the turbine and compressor. This means that almost any conceivable type of maintenance work, with the possible exception of a complete reblading of the stationary compressor and turbine sections, can be done without removing the turbine from the locomotive.

The rated speed of the turbine is 4,860 rpm, and it normally idles at 80% of this speed, or 3,888 rpm. The minimum sustaining speed is 2,680 rpm (55% of full speed) and turbine cranking requires 450 hp. Normal air requirement is 150,000 cubic feet per minute, and the maximum exhaust temperature is 860 F.

General Description of Locomotive

The complete locomotive consists of three parts: the auxiliary power unit, the maxin power unit, and the fuel tender, as shown in Fig. 7. The auxiliary and main power units together have an over-all length of 132 feet, 6 inches. The cab of each unit is carried on two 3-axle trucks, with all axles motored. The truck is of the fully equalized design, and is fitted with a bolster floating on a rubber pad to isolate it from the truck frame. Structure of the cab follows the same general design as that commonly used in diesel-electric locomotives, although the section comprising the entire length of the turbine and generators is removable for maintenance purposes.

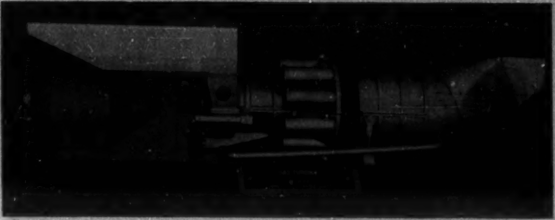


Fig. 5. Model of the gas-turbine-electric power plant for the 8,500-hp locomotive

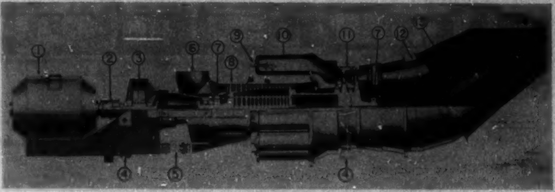


Fig. 6. Cross section of simple-cycle single-shaft 2-bearing gas turbine with reduction gear and generators for 8,500-hp gas-turbine-electric locomotive

1—Traction generator
2—Generator coupling
3—Gear box
4—Support strut
5—Connecting link
6—Air inlet
7—Main bearings
8—Axial flow compressor
9—Fuel nozzle
10—Combustion chamber
11—Two-stage turbine
12—Eductor
13—Tail cone

The auxiliary power unit contains the engineman's cab and an auxiliary diesel engine, and is capable of operating independently for light movements. Diesel fuel is carried in a 2,500-gallon tank located under the platform of this unit. A Cooper-Bessemer 6-cylinder diesel engine supplies power for the turbine accessories and other auxiliaries, and operates continuously while the locomotive is in service. On the forward end of this engine is mounted a conventional locomotive-type d-c generator with a small battery-charging generator driven through an overhung gear box. This gear box also provides a power take-off for the traction motor blower. Air from this blower is carried by ducts in the main sills to the motors in the two trucks.

A radiator cooling fan is driven from the rear end of the diesel engine through a gear box and eddy-current clutch. This gear box also drives the turbine water-coolant pump.

Two air brake compressors are provided. One is driven from the gear box that also drives the radiator fan, and the other is driven from a shaft extension on the generator at the forward end of the diesel engine. Both compressors are equipped with a disconnect feature which permits them to be uncoupled from the engine in case of compressor failure.

Auxiliary control is housed in a compartment in the conventional location immediately behind the engineman's cab. Propulsion control for this unit is located in a compartment near the rear of the cab.

The second, or main power unit, contains the power plant and associated equipment. It is not capable of operating unless coupled to the auxiliary power unit. In the forward end of this unit is a control section housing two propulsion control groups, the main generator excitation group, and the turbine starting control.

Immediately after this control area is

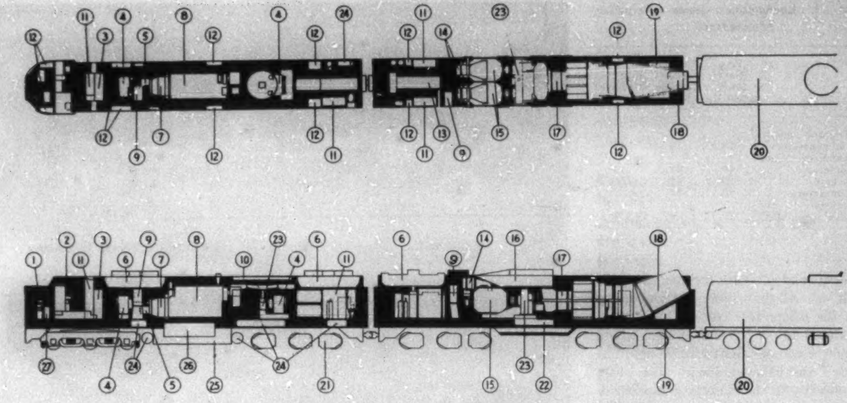


Fig. 7. Arrangement of equipment in the 8,500-hp gas-turbine locomotive

1—Air-brake equipment
2—Operator's cab
3—Auxiliary control
4—Compressor
5—D-c auxiliary generator
6—Braking resistor
7—D-c generator
8—Auxiliary diesel engine
9—Traction-motor blower
10—Radiator
11—Propulsion control
12—Sand box
13—Excitation control
14—Generator blower
15—Traction generators
16—Turbine air intake
17—Gas turbine
18—Turbine exhaust
19—Power-plant equipment
20—Gas-turbine fuel tender (furnished by customer)
21—Traction motors
22—Lubricating-oil tank
23—Gear box
24—Air reservoirs
25—Diesel fuel tank
26—Battery box
27—Train control

a separate compartment in which the mechanically driven traction motor blower is located. This supplies air to the motors on the two trucks of this unit. The generator blowers also take their air from this compartment.

The remainder of this unit is devoted to the main power plant: first, the traction generators with gear-mounted exciters and blowers; then the main reduction gear unit; finally, the turbine itself.

Connecting the main and auxiliary power units are 20 power-cable jumpers, four multiple-unit control jumpers, two coolant-water hoses, a diesel-fuel supply line for turbine starting, and the necessary air brake hoses.

Previous locomotives made wide use of time-delay relays to insure proper turbine starting sequence. In order to prevent serious consequences if these relays should operate out of sequence, it was necessary to interlock the various contactors involved. Experience has shown that these time-delay relays are subject to tampering and occasionally become maladjusted. Also, the interlocking necessary to insure safe operation added to the complexity of the locomotive. An air-driven controller, Fig. 8, was developed for turbine starting on the 8,500-hp locomotives which substitutes mechanical interlocking in the form of cams for the electric interlocking previously required and reduces the number of time-delay relays to two.

Electric Propulsion Equipment

The turbine shaft speed is reduced by means of gearing to 1,050 rpm to drive four traction generators of a type developed for use on diesel-electric locomotives. The generators are mounted, two per set, in a back-to-back arrangement with bearings at the outboard or commutator ends only. Two of these generator sets mounted side by side occupy the entire width of the locomotive cab, and no aisle is provided past them.

Unusual precautions have been taken to insure a supply of clean ventilating air for the traction generators. Each generator set drives its own air-cleaning-type blower through a gear boxes and power take-off at the forward end of the set. Air from the blower is delivered to the center of the generator set where it divides to the two machines. It passes first over the rear-end windings, then over the exciting field coils and armature, and is finally discharged from the commutator end of the generator. The advantages to be derived from this arrangement are as follows:

1. Cleaner machines, since the ventilating air is taken directly from outside the cab where a minimum of oil-vapor contamination exists.

2. More uniform cooling of the generators, since the rear-end windings, which normally reach the highest operating temperatures, receive the coolest air.

3. Better insulation life, partly because of better cleaning of the ventilating air, and partly because the direction of air flow will blow carbon dust out of the generator instead of into it.

Each of the four generators supplies power to the three traction motors on one truck. These are connected permanently in parallel and have one step of field shunting. Each group of motors with its associated control equipment, excitation equipment, and generator is completely isolated from the others, so it is possible to cut out one generator-motor combination without affecting the other three.

A number of combinations of motors and generators was considered in designing this locomotive The one chosen seemed to offer the best combination of reliability and standardization of parts.

The traction motors are the standard *GE-752* model widely used in diesel-electric, gas-turbine-electric, and rectifier-

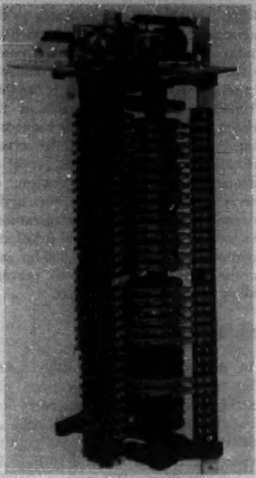


Fig. 8. Air-driven cam-type controller for turbine starting

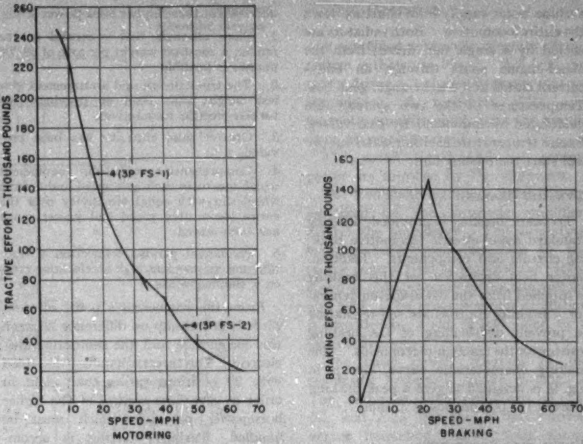


Fig. 9. Locomotive characteristics in motoring and dynamic braking for the 8,500-hp gas-turbine-electric locomotive

type electric locomotives, and incorporate the latest design improvements. They are geared with a 74-to-18 ratio to axles mounting 40-inch wheels. This gives a maximum locomotive speed of 65 mph, and produces the speed-tractive effort characteristic shown in Fig. 9.

Auxiliaries

Fuel Tender

As shown in Fig. 7, all main bunker *C* fuel is carried in a tender designed and built by the railroad. Fuel is pumped into the tender at an average temperature of approximately 180 F. For keeping the fuel warm, chief reliance is placed on the insulation of the tender tank—4 inches of resin-bound glass wool. The design of the fuel tender is such that when fuel is taken aboard at 180 F, it is possible to operate for 45 hours in an average ambient air temperature of −40 F with no supplementary heat before the minimum permissible fuel temperature is reached. In addition, two methods are provided for adding supplementary heat to keep the fuel warm indefinitely while the locomotive is operating. The first is by means of electric heaters supplied with power from the diesel-driver auxiliary generator. This heating is thermostatically controlled and is available whenever the auxiliary generator is not busy with one of its other functions. The second method is by means of a recirculating fuel line to the tender. The recirculated fuel is heated from the atomizing air supply. Since the atomizing air is bled from the main turbine compressor, this heat is available only when the turbine is operating.

Turbine Auxiliaries

The turbine governor is of the shaft-driven hydraulic fly-ball type very similar to those used on conventional diesel-electric locomotives. It is provided with a load-regulating function operated by the turbine-exhaust temperature control which serves to regulate the electric load on the generators to a level permitting safe turbine-exhaust temperature.

A fuel-metering pump of greatly improved design supplies equal amounts of fuel to all ten combustion chambers. The design of this pump resulted from experience on the existing locomotives, and has been proved in service operation. Actually, it consists of ten separate positive-displacement gear-type pumps, each feeding the fuel nozzle in one combustion chamber. A large positive-displacement pump, in series with the ten metering pumps, furnishes most of the pressure for proper fuel atomizing. The pressure drop across each of the metering pumps is only enough to adjust the pressure to the nozzle to the level necessary to secure equal distribution of fuel to all combustion chambers. This type of fuel pump allows the use of heavier fuel than the piston-type pump formerly employed. The pump is driven by hydraulic motor whose speed, in turn, is regulated by the turbine governor.

A summary of the distribution of auxiliary loads on the locomotive is given in Table II.

Cooling System

Cooling for the diesel-engine jacket water and the turbine lubricating oil is provided by a conventional water-to-air radiator located in the auxiliary unit. The water systems for the turbine and the diesel engine, however, are kept entirely separate. This was done for two reasons: first, to prevent any accidental water leakage in the interconnecting hoses or piping from shutting down both the diesel engine and the turbine; second, to prevent any accidental overheating of the

Table II. Distribution of Auxiliary Loads on the 8,500-Hp Gas-Turbine-Electric Locomotive

Diesel Engine	Main Turbine
Turbine cranking	Turbine auxiliaries
Locomotive hostling (three traction motors)	Traction motor blower on main power unit
Traction motor blower on auxiliary power unit	Traction generator blowers
Electric heat for fuel tender (70%)	Recirculated fuel heat for tender (30%)
Battery charging	
Excitation for dynamic braking	
Coolant water pump	
Lubricating oil and fuel pumps	
Radiator fan	
Air compressors (two)	

turbine water supply from shutting down the entire locomotive. Both radiators are cooled by a single fan, driven from the diesel-engine shaft through an eddy-current clutch and a right-angle gear box. Temperatures of the two systems are modulated by individual by-pass valves. Excess temperature in either water supply will start the cooling fan.

Dynamic Braking

Dynamic braking is provided, utilizing standard diesel-electric locomotive braking circuits and components. Traction-motor field excitation for this braking is supplied from the diesel-driven generator. Braking controls are self-regulating to prevent overloading of the braking resistors or the traction-motor fields. The braking characteristic curve, shown in Fig. 9, is designed to give a peak braking effort of 147,000 pounds at 22 mph.

Wheel-Slip Protection

Added attention has been paid in the design of this locomotive to the control of wheel slip. At the nominal horsepower rating of the turbine, each traction motor is expected to handle approximately 708 hp. This is compared with 600 hp per motor for the first series of gas-turbine-electric locomotives and 425 hp for typical diesel-electric locomotives. Five steps have been taken to maintain good adhesion at these higher horsepower levels:

1. Since all main fuel is carried on the tender, a constant weight per axle of 68,000 pounds is possible.
2. The truck design and arrangement give less weight shift than on previous gas-turbine-electric locomotives.
3. Greater sand capacity has been provided.
4. Comprehensive wheel-slip equipment which measures each axle speed and detects wheel slip with equal sensitivity over the entire locomotive speed and power range has been added.
5. Permanent parallel connection of traction motors prevents high acceleration rates on a slipping wheel.

From the engineman's point of view, there is practically no difference between this locomotive and the standard diesel-electric. The throttle handle is provided with 20 positions rather than eight in order to give finer control of the higher horsepower per axle which must be handled. Turbine starting is accomplished by simply pushing a button on the engineman's control panel. Once this button is depressed, the turbine starting sequence takes over automatically and follows through the complete cycle lasting approximately 9 minutes. As soon as the cycle is complete this fact is indicated to the engineman.

In addition to the standard air-brake gauges and ammeters, the engineman has one additional meter marked "Notching Guide." In general, it indicates to the engineman the condition of the turbine; that is whether there is a reserve of power in the turbine or whether he is demanding all the power of which it is capable. This meter was found valuable on the existing locomotives, and has been continued in the new design. Under certain extreme operating conditions, such as high-temperature tunnel operation, or during turbine acceleration, or under extremely high ambient temperatures, this indication has been useful in helping enginemen on the present 4,500-hp gas-turbine-electric locomotives to prevent stalling the turbine power plant.

Conclusions

In designing this locomotive, every effort has been made to use parts and practices already developed and proved for diesel-electric locomotive service. A great deal of attention has been paid to keeping the number of these parts to an absolute minimum. The turbine itself is, of course, new; but, it is of a design backed by 8 years of operating experience, plus considerable advance in the design and construction of gas turbines in general. This combustion should provide a locomotive which, properly applied, will set new levels of performance at costs competitive with other existing forms of motive power.

Vol.
Pt. 2
1957

Design and Analysis of the Motor-Alternator Speed Regulator for the New Haven Speed-Merchant Locomotive

S. W. McELHENNY
ASSOCIATE MEMBER AIEE

THE Speed-Merchant locomotive is a dual-power high-speed lightweight locomotive. Traction and auxiliary power is supplied either by its diesel engine or from the d-c third-rail source.

A 350-horsepower 3-phase 440-volt alternator in the locomotive provides power for train heating, air conditioning, lighting, and water heating and cooling.

During diesel operation, the alternator is driven at constant speed by the diesel engine, regardless of throttle notch position. During electrified zone operation, however, the diesel engine is necessarily inoperative and the alternator is driven by a d-c motor operating from the third-rail source. The speed of this motor must be maintained nearly constant, in order to maintain alternator frequency within the limits necessary to insure proper operation of the equipment in the cars.

Because the supply voltage varies from 450 to 725 volts and the a-c load varies from 50 to 350 horsepower, a regulator is needed to control the motor speed. Moreover, the supply voltage variation is not necessarily gradual, but can be a double-step function from 450 to 0 to 725 volts, or conversely, as the locomotive passes over a third-rail gap. The zero voltage time is determined by train speed and length of the third-rail gap.

The system was designed using the steady-state sinusoidal frequency-attenuation criteria. The severe operating disturbances made a transient analysis desirable and an electronic differential analyzer was used. These two techniques go hand in hand. The chief value of the computer lies in its ability to produce quantitative results quickly for all operating conditions, especially transients.

Paper 57-15, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 19, 1956; made available for printing November 7, 1956.

S. W. McElhenny is with the General Electric Company, Erie, Pa.

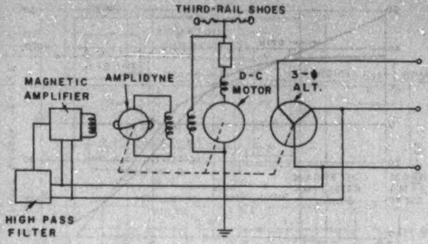

THIRD-RAIL SHOES
MAGNETIC AMPLIFIER
AMPLIDYNE
D-C MOTOR
3-ϕ ALT.
HIGH PASS FILTER

Fig. 1. Simplified schematic diagram

System Operation

A simplified schematic diagram of the system is shown in Fig. 1. A 3-field d-c motor drives a 3-phase alternator and an amplidyne. A high-pass filter and magnetic amplifier work with the amplidyne to regulate speed.

The speed signal and a-c control power are supplied by the alternator.

The nonlinear high-pass filter output, amplified by the magnetic amplifier, excites the amplidyne. The amplidyne supplies buck-boost excitation to the motor control field to regulate the motor speed as changes occur in the line voltage and load. The primary motor excitation is provided by a shunt field aided by a cumulative series field.

The basic circuit diagram of the regulating system is shown in Fig. 2. The high-pass filter composed of *C1*, *C2*, and *RE2* has the output current versus frequency characteristic shown in Fig. 3(A).

The filter output current coupled by matching transformer *T2* is rectified by bridge *RT1* and is subtracted from a reference current supplied by the regulated d-c source to provide a net signal through the *F1-F2* winding. This signal will bias off the magnetic amplifier at frequencies below approximately 53 cps (cycles per second). The d-c gain of the

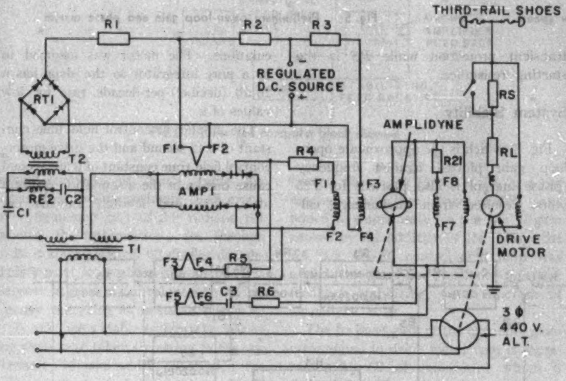

R1
R2
R3
REGULATED D.C. SOURCE
-
+
RT1
T2
F1
F2
RE2
C2
C1
AMP1
R4
F1
F2
F3
F4
T1
F3
F4
R5
F5
F6
C3
R6
AMPLIDYNE
R21
F8
F7
THIRD-RAIL SHOES
RS
RL
DRIVE MOTOR
3 ϕ 440 V. ALT.

Fig. 2. Basic circuit diagram

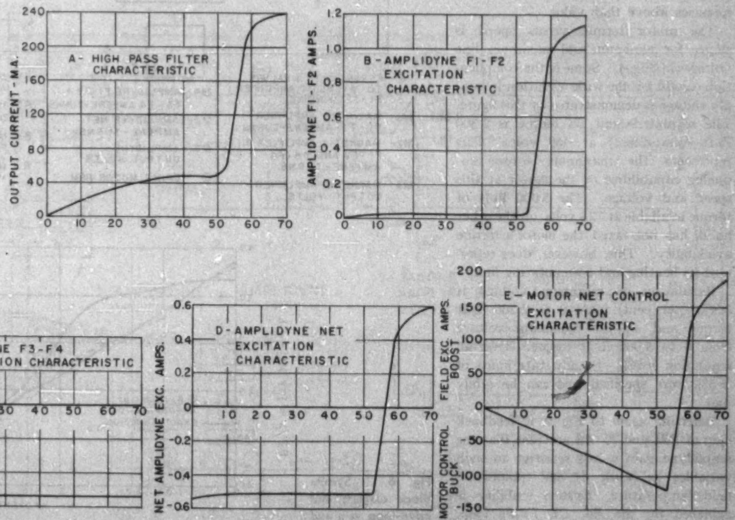


Fig. 3. Performance characteristics plotted against frequency in cps

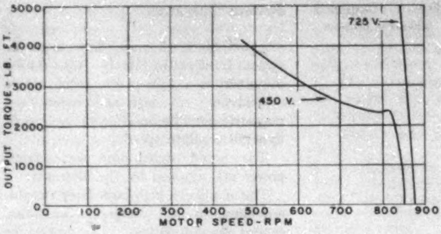


Fig. 4. Output torque versus motor speed

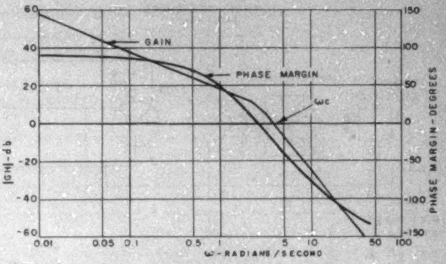


Fig. 5. Preliminary open-loop gain and phase margin

magnetic amplifier is controlled by the negative feedback loop composed of *R5* and magnetic amplifier winding *F3-F4*. Thus, the signal in the amplidyne *F1-F2* winding, shown in Fig. 3(B), has a characteristic similar to that of the high-pass filter.

The regulated d-c source also provides a constant signal through the amplidyne *F3-F4* winding, regardless of frequency. This is seen in Fig. 3(C).

The net amplidyne excitation is the difference between the *F1-F2* and *F3-F4* signals and is plotted against frequency in Fig. 3(D). The resulting motor-control field excitation characteristic is shown in Fig. 3(E). As can be seen, the motor-control field excitation bucks the main excitation at frequencies below the nominal of 56 cps and boosts it at frequencies above that value.

The motor torque versus speed is shown for minimum and maximum line voltages in Fig. 4. Some of the complication caused by the wide variation in supply voltage is demonstrated by this figure. The regulated limit on torque is 2,300 lb-ft (pound-feet) at 450 volts. This represents the maximum torque-producing capabilities of the motor at this speed and voltage. The 5,000 lb-ft of torque available at 725 volts, on the other hand, has not taxed the motor's torque availability. This, however, does represent its heating and commutating limit.

Regulation at minimum voltage is 3.8% (per cent) from 0 to 2,200 lb-ft torque and 4% at maximum voltage from 0 to 5,000 lb-ft torque. Over-all regulation is 7%. Steady-state limits of ±5% were specified and can be safely met.

Referring again to Fig. 2, a feedback loop of *R21* and *F7-F8* provides that the amplidyne gain is less sensitive to such variables as brush fit and quadrature field temperature. System stability is achieved by the *R6, C3, F5-F6* loop. *RL* is a permanent series resistance for transient protection while *RS* is the starting resistance.

System Stability

Fig. 5, which is the approximate open-loop gain plotted against frequency (phase margin is also plotted) for 725 volts, resulted from preliminary calculations. The motor was assumed to be a pure integrator so the slope has a 20-db (decibel)-per-decade rate for low values of ω.

The amplidyne-control field time constant of 0.42 second and the drive-motor-control field time constant of 0.306 second cause breaks in the attenuation diagram at ω=2.38 and ω=3.27 radians per

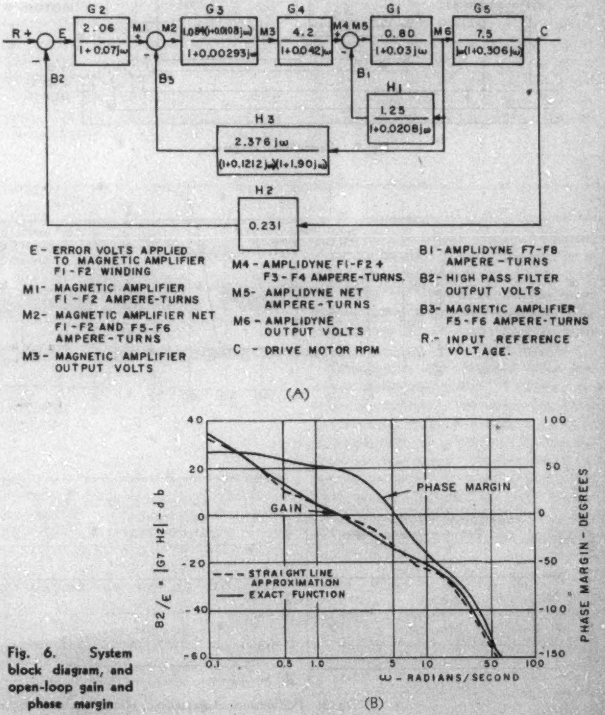

R +
E
G2
2.06
1+0.07jω
M1
M2
G3
1.084(1+0.0108jω)
1+0.00293jω
M3
G4
4.2
1+0.042jω
M4 M5
G1
0.80
1+0.03jω
M6
G5
7.5
jω(1+0.306jω)
C
B2
B3
B1
H1
1.25
1+0.0208jω
H3
2.376jω
(1+0.1212jω)(1+1.90jω)
H2
0.231
E - ERROR VOLTS APPLIED TO MAGNETIC AMPLIFIER F1-F2 WINDING
M1 - MAGNETIC AMPLIFIER F1-F2 AMPERE-TURNS
M2 - MAGNETIC AMPLIFIER NET F1-F2 AND F5-F6 AMPERE-TURNS
M3 - MAGNETIC AMPLIFIER OUTPUT VOLTS
M4 - AMPLIDYNE F1-F2 + F3-F4 AMPERE-TURNS.
M5 - AMPLIDYNE NET AMPERE-TURNS
M6 - AMPLIDYNE OUTPUT VOLTS
C - DRIVE MOTOR RPM
B1 - AMPLIDYNE F7-F8 AMPERE-TURNS
B2 - HIGH PASS FILTER OUTPUT VOLTS
B3 - MAGNETIC AMPLIFIER F5-F6 AMPERE-TURNS
R - INPUT REFERENCE VOLTAGE.
(A)
PHASE MARGIN
GAIN
STRAIGHT LINE APPROXIMATION
EXACT FUNCTION
40
20
0
-20
-40
-60
B2/E = |G7 H2| - db
100
50
0
-50
-100
-150
PHASE MARGIN - DEGREES
0.1
0.5
1.0
5
10
50
100
ω - RADIANS/SECOND
(B)

Fig. 6. System block diagram, and open-loop gain and phase margin

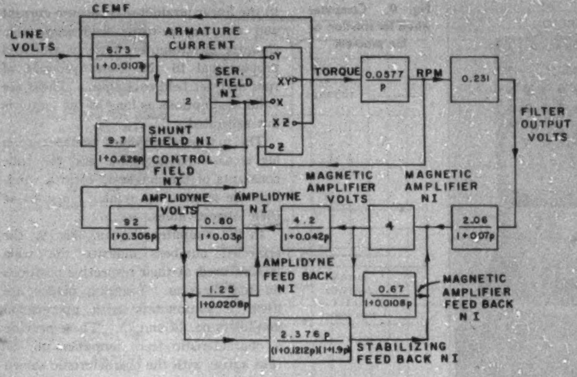

CEMF
LINE VOLTS
ARMATURE CURRENT
6.73 / 1+0.010p
SER. FIELD NI
2
Y
XY
X
X Z
Z
TORQUE
0.0577 / p
RPM
0.231
FILTER OUTPUT VOLTS
9.7 / 1+0.626p
SHUNT FIELD NI
CONTROL FIELD NI
MAGNETIC AMPLIFIER VOLTS
MAGNETIC AMPLIFIER NI
AMPLIDYNE VOLTS
AMPLIDYNE NI
92 / 1+0.306p
0.80 / 1+0.03p
4.2 / 1+0.042p
4
2.06 / 1+0.07p
AMPLIDYNE FEED BACK NI
1.25 / 1+0.0208p
0.67 / 1+0.0108p
MAGNETIC AMPLIFIER FEED BACK NI
2.376 p / (1+0.1212p)(1+1.9p)
STABILIZING FEED BACK NI

Fig. 7. Unscaled computer block diagram

second respectively. Since these are lower than the approximate 0 db crossover frequency (ω_c) of 3.9 radians per second, the attenuation rate through 0 db is approximately 60 db per decade. This results in a phase margin of −25 degrees at the 0 decibel crossover frequency indicating an unstable system.

To produce a stable system, the following steps were taken (referring to Fig. 2). Resistor *R4* was added to shorten the time constant of the amplidyne *F1-F2* control field by a factor of 10, and the magnetic amplifier stage *AMP1* was added to compensate for the loss of gain caused by the addition of *R4*. Also, the *R6*, *C3*, *F5-F6* control field loop was added to provide frequency-sensitive feedback stabilizing, as a function of amplidyne volts.

The forward gain of the elements from error signal to drive motor rpm is represented by *G7* of equation 1 which is defined by Fig. 6(A). The open-loop system gain *G7H2* is plotted against frequency (both magnitude and phase margin) in Fig. 6(B) and is defined by equation 2 and Fig. 6(A).

$$G7=\frac{G1G2G3G4G5}{1+G1(H1+G3G4H3)}=\frac{C}{E} \qquad (1)$$

and

$$G7H2=\frac{G1G2G3G4G5H2}{1+G1(H1+G3G4H3)}=\frac{B2}{E} \qquad (2)$$

The open-loop gain is maintained at its original level for low values of ω. This provides satisfactory steady-state regulation. The frequency attenuation diagram is modified between $\omega=0.24$ and $\omega=19$ radians per second primarily by the addition of the *R6*, *C3*, *F5-F6* loop. The constants in the stabilizing loop cause the increase in attenuation rate at $\omega=0.24$ as well as the decrease in rate at $\omega=0.526$. This allows the attenuation curve to pass through unity gain (0 db) at approximately a 20-db-per-decade rate. This provides for a phase margin of +48 degrees at the unity gain frequency of $\omega=1.65$ radians per second indicating that the system is stable.

Computer Analysis and Test Results

The computer block diagram is shown in Fig. 7. The symbol p is the differential operator.

Torque is assumed to be proportional

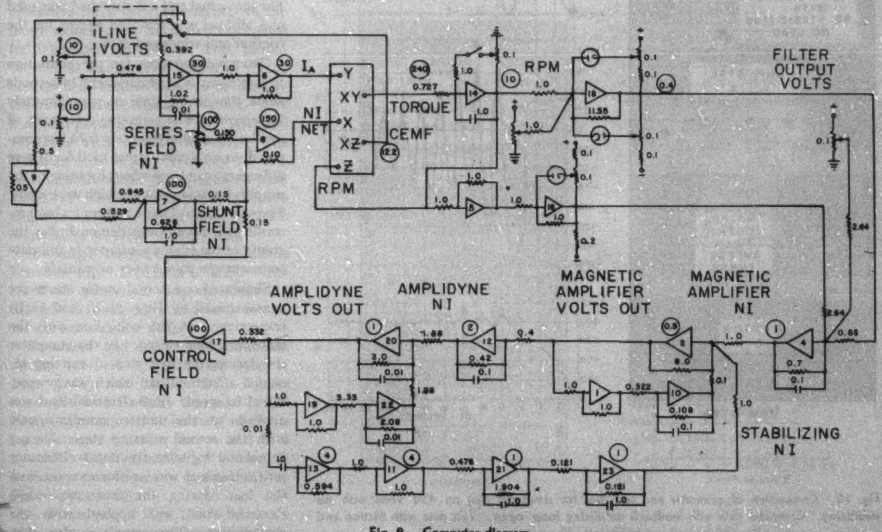

LINE VOLTS
SERIES FIELD NI
SHUNT FIELD NI
I_A
NI NET
Y
XY
X
XZ
Z
TORQUE
CEMF
RPM
RPM
FILTER OUTPUT VOLTS
CONTROL FIELD NI
AMPLIDYNE VOLTS OUT
AMPLIDYNE NI
MAGNETIC AMPLIFIER VOLTS OUT
MAGNETIC AMPLIFIER NI
STABILIZING NI

Fig. 8. Computer diagram

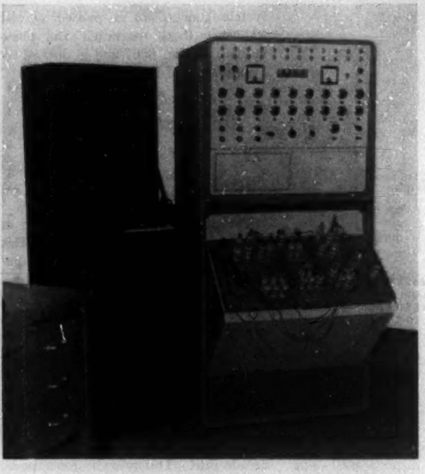


Fig. 9. Computer wired for solution of the problem

to the linear product of armature current and net ampere-turns field excitation. Counter electromotive force is assumed proportional to the linear product of rpm and net field excitation. These are valid assumptions as long as the motor is not saturated.

The magnetic amplifier transportation lag is assumed negligible and the time constants of the magnetic control windings are assumed to remain constant at their longest values.

In the computer diagram, Fig. 8, the encircled numbers indicate the scale factors used at their respective positions on the diagram. Limiting diodes are used in conjunction with operational amplifiers no. *16* and *18*. These provide a characteristic from amplifier no. *17* that agrees with the characteristic shown in Fig. 3(E).

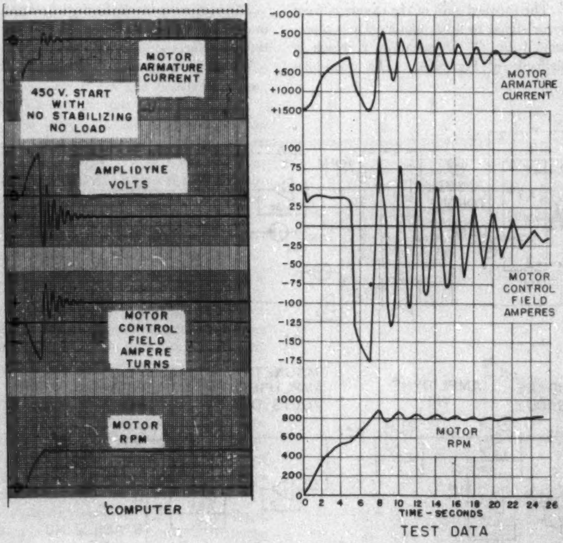
The computer wired for the problem is shown in Fig. 9.

A comparison of the computer recordings and reduced oscillographic data is shown in Figs. 10 through 13.

For the computer simulation, 2-step starts were used with the exception of Fig. 10(A), which demonstrates a simulated start on 450 volts line with the feedback stabilizing loop open. All friction and electrical load were purposely removed for this recording, as this was determined to be the most unstable condition. Fig. 10(B) is a plot of the oscillographic data for the same transient. For the actual test a 2-step start was used and all load was removed except for the friction and windage losses.

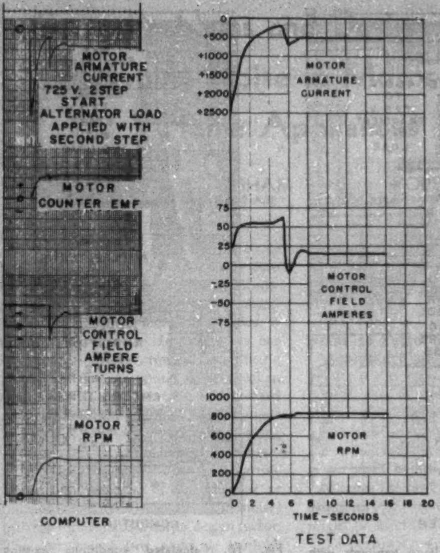
The predicted time for the oscillation to damp out is approximately 18 seconds while the actual time is approximately 26 seconds. The computer frequency of oscillation is approximately 0.75 cps compared to an actual value of 0.5. These differences can be resolved by some of the simplifying assumptions which were made. The same general start pattern on both recordings is noted, and demonstrates the ability of the analog computer to simulate transient conditions very accurately.

Simulated and actual motor starts are demonstrated by Figs. 11(A) and 11(B) respectively, at 725 volts line with the stabilizing loop closed. In the computer simulation, the first step of starting included a friction load which was proportional to speed. Full electrical load was applied to the motor simultaneously with the second starting step. During actual starting tests, the motor-alternator set friction and windage losses comprised the load during the first step while electrical load was applied with the second step. The actual counter elec-



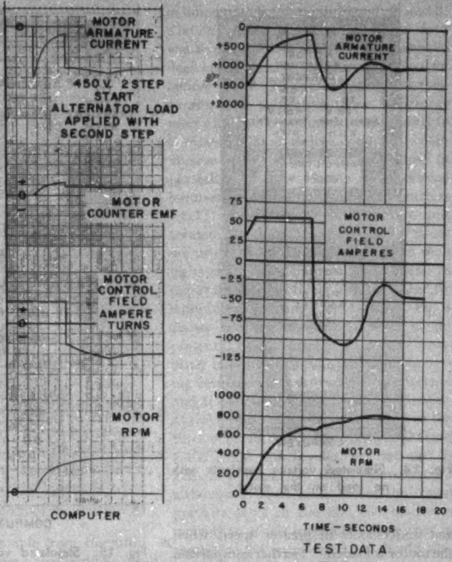
(A) (B)

Fig. 10. Comparison of computer and test data for simulated start on 450 volts with no stabilizing. Computer data with feedback stabilizing loop open. Test data with friction and windage load



(A) (B)

Fig. 11. Simulated and actual motor starts, 725 volts, 2-step start, alternator load applied with second step



(A) (B)

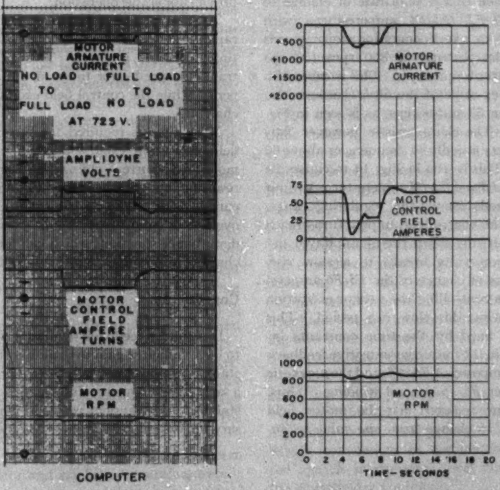
Fig. 12. Simulated and actual motor starts, 450 volts, 2-step start, alternator load applied with second step

tromotive force per rpm constant of the motor exceeded the calculated value used for computer simulation. This caused the actual steady-state values of motor control field ampere-turns to be slightly different from the simulated values. The transient pattern in both Figs. 11(A) and 11(B) is very similar, however. The transient time is approximately 7.5 seconds in both cases. The simulated peak value of armature current is 2,480 amperes while the actual value is 2,300 amperes, or an error of only 7.8%.

A similar comparison can be made between Figs. 12(A) and 12(B). The transient period is approximately 18 seconds in both cases and both sets of curves have similar shapes.

Simulated and actual load transients during 725-volt operation are seen in Figs. 13(A) and 13(B). Both figures show the load increase transient period lasting approximately 2 seconds, while the load decrease transient decays in 3 to 4 seconds. Distinct similarity is seen in both a quantitative and qualitative analysis of the two figures.

A comparison of Figs. 14 and 15 shows that voltage transients cause overshoots



(A) (B)

Fig. 13. Simulated and actual load transients at 725 volts

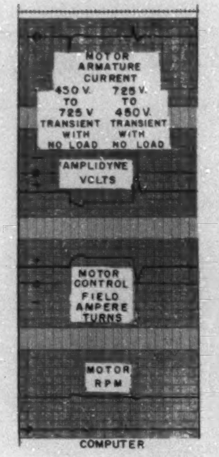


Fig. 14. Simulated voltage transients with no load on the motor

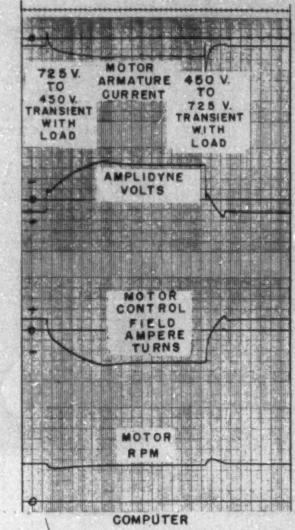


Fig. 15. Simulated voltage transients with load on the motor

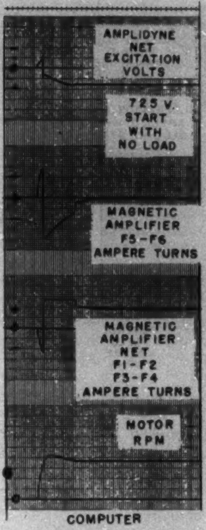


Fig. 16. Calculated conditions existing during start at 725 volts with no load on the motor

and undershoots at greater speed when the motor is loaded. Further comparison shows that the same voltage transient demands a greater change of motor-control field excitation when the motor is operating under load.

The cause and amount of transient speed error is best explained by Fig. 16. It is seen that a high rate of change of net *F1-F2*, *F3-F4* ampere-turns with respect to speed occurs for motor speeds near the nominal of 850 rpm (56 cps alternator frequency). This is caused by the steep slope of the filter characteristic between 53 and 60 cps, as is seen in Fig. 3(A). The characteristic, however, flattens very rapidly at frequencies above 60 cps. This is seen in Fig. 16 to cause the net *F1-F2* and *F3-F4* ampere-turns to limit for speeds above the regulating range. After the controlling ampere-turns reach the limit, the sharp rise of amplidyne net excitation volts breaks to a slow rise. The rate of change of the *F5-F6* ampere-turns is controlling the rate of excitation rise during this slow rise period. This rate is fixed by the time constants involved in the frequency-sensitive feedback loop composed of *R6*, *C3*, and the magnetic amplifier *F5-F6* control winding. Thus, the time required for the motor field current to change from one value to another is controlled primarily by the frequency-sensitive feedback loop time constants which were made purposely long, to provide for stable operation.

The motor armature current can change rapidly, however, because the armature time constant is short. This allows a transient speed error to occur during the period while the control field excitation is changing.

More time is required for the control field excitation excursion when a large motor field current change is demanded, resulting in a greater transient motor field current error. Thus, greater motor speed overshoots and undershoots occur for disturbance functions which call for large changes in motor control field excitation.

Conclusions

Designing a feedback regulating system to have sufficient gain to maintain steady-state limits does not necessarily produce a satisfactory system. Information about stability and transient response is necessary in order to determine whether the system will meet the specified conditions.

Use was made of the frequency attenuation diagram method for the selection of circuit constants. This made it possible to synthesize a stable system having sufficient gain to regulate within specified steady-state limits.

The electronic differential analyzer made it possible quickly to analyze and evaluate the system response to all expected transients under all known operating conditions before physical construction of the system was started. It was also possible to determine the effects of component tolerances and gain changes with great facility.

Actual system tests proved the validity of the data which were secured by using the analyzer.

When using the computer, unexpected phenomena are discovered before the system is constructed. Thus, the long hours of tedious calculation and field testing are largely eliminated. Without question, the analog computer is an essential tool of the modern system designer.

Some Problems and Design Factors Concerning Noise in Motor-Driven Home Appliances

E. R. CUNNINGHAM
ASSOCIATE MEMBER AIEE

G. L. WOLFERT
ASSOCIATE MEMBER AIEE

DURING RECENT YEARS, the ratio of material to labor cost of mass-produced products has been increasing, mainly owing to the more effective use of automatic processes. In turn, this has stimulated efforts toward producing greater output per pound of material in such items as small motors, and toward greater utility per pound in such products as home appliances. This trend, along with the shifting of appliances from basement to kitchen or first-floor utility areas, brings the reduction of objectionable or undesirable sound (referred to as noise) into the forefront as a major design factor. The purpose of this paper is twofold. First, the over-all appliance noise problem in general terms will be discussed and, second, the problem of reducing that portion of noise contributed by the motor will be discussed specifically.

Nomenclature

$a = \frac{\text{fundamental conductors on } s\text{-phase}}{\text{fundamental conductors on } m\text{-phase}}$

c = existing damping constant

c_c = critical damping constant

f = line frequency

I_m = effective value of m-phase current

I_s = effective value of s-phase current

K = spring constant

n = actual motor speed

n_s = synchronous speed

P = pairs of stator poles in motor

R = acoustical room constant $= \frac{\alpha S}{1-\alpha}$

R_b = apparent resistance to m-phase backward field

R_f = apparent resistance to m-phase forward field

S = surface area of a room

S_r = number of rotor teeth

s = slip $= \frac{n_s - n}{n_s}$

X_b = apparent reactance to m-phase backward field

X_f = apparent reactance to m-phase forward field

α = average absorption coefficient of room surfaces

ϕ = angle by which I_s leads I_m

Noise as Factor in Customer Acceptance and Satisfaction

Objection to noise divides into two, at least partially separable, parts: that which can be termed psychological and that which falls into the category of loudness. Although the prime purpose here is to discuss the problem of loudness, a few examples of psychological annoyance might be in order.

To borrow an example from the automotive industry, many individuals accept, with some feeling of pleasure, the low-frequency exhaust roar accompanying acceleration (suggesting power), but are quite perturbed by a slight ticking in the motor (suggesting improper functioning) or similarly by an illusive rattle under the dashboard. The sound level of one can be measured quite easily with common sound-measuring instruments; the other is so low in relation to the accepted background noise that the use of a narrow-band analyzer and other specialized equipment are necessary for detection. Similarly, obtaining the pleasing "clunk" with the closing of car doors instead of a "clink" has been a definite and well-recognized problem in the automotive industry.

Similar problems exist in the field of home appliances where various and sundry noises, such as relay buzz, intermittent pump and gear noises, intermittent clinking of buttons on the revolving drum of a dryer, etc., can be quite annoying to some people when the appliance is located in close proximity to the living quarters.

Generally, it is often difficult to determine the factors related to annoyance since such items as previous conditioning of individuals is involved as well as the character of the noise.

Of perhaps more immediate concern to motor and appliance manufacturers is the actual loudness of the sound produced by their products. Examples of noise measurements taken of appliance noise in homes can be found in Figs. 1, 2, and 3. The method used to obtain these data was to measure the sound pressure level in decibels (objective measurement) on an octave basis. This information was then expressed in subjective terms to present more effectively a numeric proportional to loudness. To describe loudness, conversion of the octave data to "phons" (a partially subjective means of describing loudness in decibels, and termed "loudness level") or to "sones" (a subjective unit termed "loudness") may be made. Sones are not expressed in decibels, but in such units that a doubling of the loudness, as determined subjectively, is represented by a doubling in the number of sones. Fig. 4 shows the relationships between sound pressure level, phons, and sones.[1] Another important consideration in evaluating noise data is that of speech interference and telephone usability. The arithmetic average of the sound pressure level in the three octaves, 600/1,200, 1,200/2,400, and 2,400/4,800 cps (cycles per second) is generally used for this determination; although in such cases where the sound pressure level in the 300/600 octave is more than 10 db (decibels) above that in the 600/1,200 cps octave, it should also be included in the arithmetic average. This arithmetic average is often referred to as the "speech interference level."

Although the level at which normal conversation can be easily carried on in the home depends on many factors, such as separation distance of individuals and the room acoustics, generally whenever the speech interference level exceeds 55–60 db, normal conversation at separation distances of 3 feet or more becomes difficult, as does telephone usability.

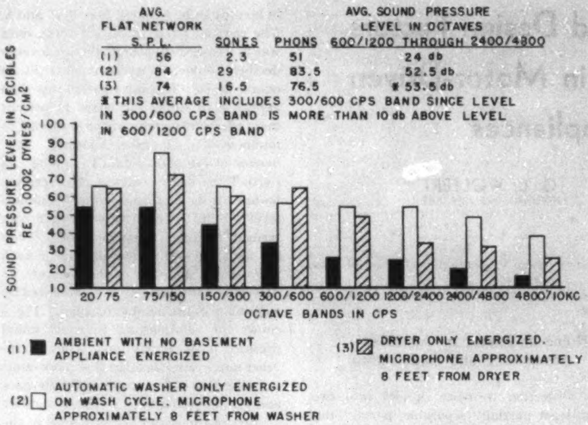
The objective measurements taken in three different homes as shown on Figs. 1, 2, and 3 have been converted to obtain loudness levels, loudness, and the speech interference levels. The appliances in location *A* (Fig. 1) were considered acceptably quiet by the householder; the automatic washer in location *B* (Fig. 2) was considered noisy by the householder involved, and the noise presented a definite telephone usability problem (the telephone was also located in the kitchen); the automatic washer in location *C* (Fig. 3) was considered to have an objectionably loud pure tone of 120-cps noise which resulted in standing waves of sound at this frequency throughout the basement and first-floor rooms. In general, the subjective conversions in conjunction with the speech interference level confirmed the householder's opinions.

There are, of course, important factors governing noise acceptance in the home

---

Paper 57-27, recommended by the AIEE Domestic and Commercial Applications Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted September 13, 1956; made available for printing December 12, 1956.

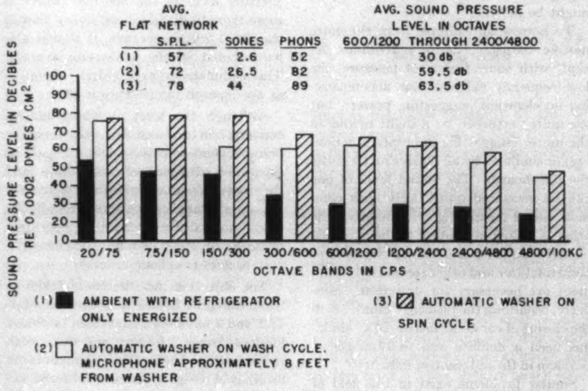
E. R. Cunningham and G. L. Wolfert are with the General Electric Company, Fort Wayne, Ind.



	AVG. FLAT NETWORK S.P.L.	SONES	PHONS	AVG. SOUND PRESSURE LEVEL IN OCTAVES 600/1200 THROUGH 2400/4800
(1)	56	2.3	51	24 db
(2)	84	29	83.5	52.5 db
(3)	74	16.5	76.5	* 53.5 db

* THIS AVERAGE INCLUDES 300/600 CPS BAND SINCE LEVEL IN 300/600 CPS BAND IS MORE THAN 10 db ABOVE LEVEL IN 600/1200 CPS BAND

Fig. 1. Location A: suburban home basement, evening hours



	AVG. FLAT NETWORK S.P.L.	SONES	PHONS	AVG. SOUND PRESSURE LEVEL IN OCTAVES 600/1200 THROUGH 2400/4800
(1)	57	2.6	52	30 db
(2)	72	26.5	82	59.5 db
(3)	78	44	89	63.5 db

Fig. 2. Location B: suburban kitchen laundry equipment located on concrete subfloor

which are not directly under the control of the manufacturers. These are:

1. Ambient or background noise level prevailing in the appliance area.
2. Acoustical properties of the room in which the appliance is located.
3. Floor construction on which the appliance is mounted.

The first of these factors is quite obviously important. Other conditions being equal, for example, an appliance will sound relatively quieter in a busy department store than in a small suburban store, since the background noise available to mask the appliance noise is generally much higher in the first case.

The acoustical properties of the room are also important. Reference to Fig. 5 shows the difference in sound pressure level which would exist as a function of distance from a nondirectional noise source in rooms having different acoustical properties. Calculations by the authors have indicated that the value of R for various kitchens and basement areas run from about 50 to 300 over the frequency range of importance. It can be expected, then, that the sound pressure level at distant points of the kitchen or basement in which the appliance is located will be of the same order of magnitude as the average level measured under free field conditions at a distance of one to two feet from the appliance. (Free field is a sound field free from reflecting surfaces.)

The type of floor construction on which the appliance is mounted can have a definite influence on resulting loudness. A wooden floor, for instance, is more flexible than a solid concrete floor, and any vibration coupled from the appliance to the floor can be reradiated from the floor and also transferred through joists and studding to other rooms where the resulting noise may be a cause for dissatisfaction. The manner in which the appliance feet contact the floor can also affect the level of noise radiated directly from the appliance. The authors have noted variations in the level of 120-cps noise, for example, of as much as 10 db, by merely moving an automatic washer over an uneven concrete floor.

These are some of the factors involved in obtaining customer acceptance. More specifically, the concern here is with the contribution of the motor to the appliance noise level and methods for reducing motor induced noise.

Nature of Motor Forces

Electromagnetic Forces

The electromagnetic forces associated with a single-phase induction motor can be broadly classified as follows:

1. Slot harmonic forces.
2. Dissymmetry harmonic forces.
3. Torque pulsation forces.
4. Solenoidal pulsation forces.

The slot harmonic forces occur because of the existence of a finite number of rotor and stator slots. The basic frequencies involved are $Sr(1-s)f/P$, $Sr(1-s)f/P+2f$, and $Sr(1-s)f/P-2f$ cps.[2] Both the pole structure and frequencies associated with these forces can be controlled by proper choice of rotor and stator slot numbers, within, of course, the limits of proper electrical performance and manufacturing ease. The magnitude of these forces increases with increasing air-gap flux density. To limit the magnitude with the increased flux densities, motor manufacturers utilize such methods as proper rotor tooth

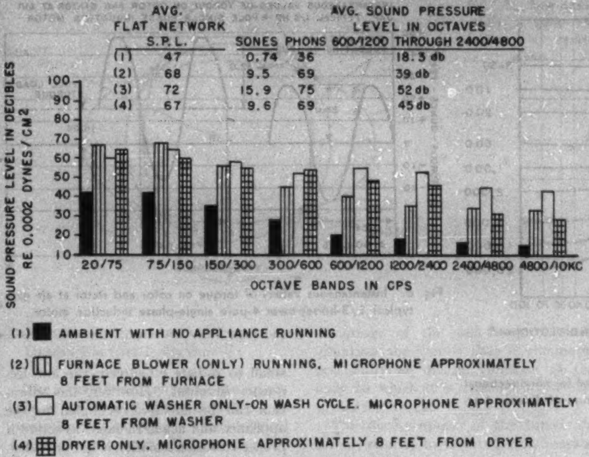


Fig. 3. Location C: basement of urban home

number, rotor tooth spiral, and closed slot rotors.

Dissymmetry forces occur because of the existence of electrical and mechanical dissymmetries. Included are such items are imperfect rotor castings, variations of bridge thickness around the rotor circumference, and off-centering of the rotor with respect to the stator bore.

Torque pulsations are inherent in a single-phase induction motor and occur at a basic frequency of $2f$ cps. The magnitude of the oscillating torque can be determined by the following equation;[3] see Fig. 6.

$$T_{\text{oscillation max}} = \sqrt{(I_m^4+a^4I_s^4+2I_m^2I_s^2a^2\cos 2\phi)\times (R_f-R_b)^2+(X_f-X_b)^2}$$

for a permanent split-phase capacitor motor or

$$T_{\text{oscillation max}} = I_m^2\sqrt{(R_f-R_b)^2+(X_f-X_b)^2}$$

for a true single-phase induction motor.

A review of these equations shows that a permanent split-phase motor can have considerably reduced torque pulsations. Thus, when the rotor bars are skewed or spiraled, a component of this force acts in the axial direction of the motor. The axial component of this force can be reduced by herringbone skewed rotor bars.

The force exerted by a magnetic field is always such as to maximize the magnetic energy and minimize the electric energy. Thus, an axial solenoidal force acts so as to reduce the air gap between rotor and stator to an effective minimum, which position is generally referred to as magnetic neutral. This force pulsates at a frequency of $2f$ cps. The magnitude of this force depends upon the air-gap flux density, relative lengths of rotor and stator, and rotor-to-stator axial positions.

There are, of course, mechanical forces, such as inertia forces due to rotor unbalance or axial forces caused by rotor to stator thrust surface wobble, which can be of concern when of sufficient magnitude. However, in most cases these are manufacturing problems which are adequately controlled.

Nature of Vibrations and Noises Produced by Motor Forces

The forces within motors which have just been described act on the various parts to cause vibration. Vibration as such is not necessarily objectionable, but it may be so if it effectively couples to the air so as to cause excessive noise. This vibration may result in sound radiated directly from the motor, or may be transmitted through the mounting supports to structural elements of the appliance from which it may be considered to be indirectly (with respect to the motor) radiated as sound. The amplitude of a sound wave depends upon the amplitude of air particle vibration which in turn depends on: 1. The magnitude and frequency of the vibrating part, and 2. the

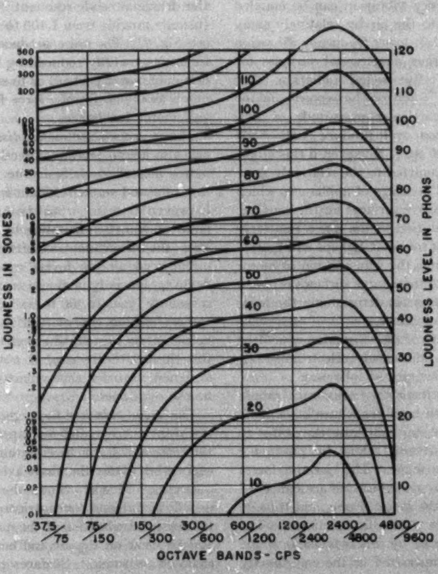


Fig. 4. Relationships between sound pressure level, phons, and sones[1]

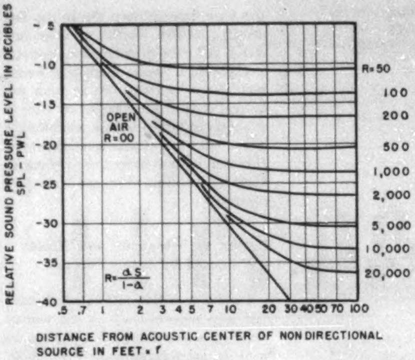


Fig. 5. Sound pressure level relative to power level for nondirectional source of different values of room constant as function of distance from source

s = total area of boundary surfaces, square feet
α = average absorption coefficient of room surfaces
SPL = sound pressure level
PWL = power level
R = acoustical room constant

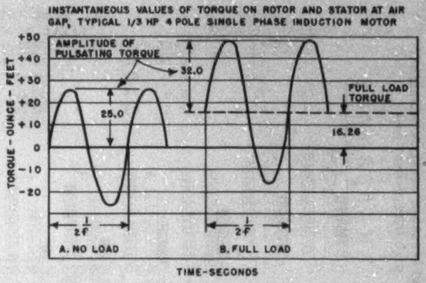


Fig. 6. Instantaneous values of torque on rotor and stator at air gap, typical 1/3-horsepower 4-pole single-phase induction motor

effective coupling of the vibration to the air.

With respect to the effective coupling, high-frequency vibration can be coupled effectively to the air by relatively small surfaces, while lower frequency vibration requires larger dimensional surfaces for equally effective sound radiation. For example, a fractional-horsepower motor having a particular magnitude of 120-cps vibration will produce very little sound at 120 cps. However, if this vibration is transmitted to the relatively large sheet-metal housings of appliances which provide effective surfaces for the coupling of such vibration to the air, a relatively high level of sound at 120 cps will be produced. Thus, the transmission of motor vibrations to these large surfaces is one of the vital factors in determining the noise level of the appliance.

The nature of the vibrations which the previously discussed electromagnetic forces produce are as follows:

The tooth frequency forces, with proper choice of slot numbers, cause little or no rotor flexure, but do cause the stator to vibrate in a radial plane by producing changes in shape. The inertia forces developed by this vibration are cancelled out within the motor frame, and thus do not produce external reactions. The noise generated by this movement of the stator is transmitted to the ear directly from the motor frame and, with most motor-mounting systems, is not coupled to the appliance. These vibrations are usually highest in amplitude at or near the frame 4 node-resonant frequencies (usually ranging from 1,100 to 1,600 cps); see Fig. 7. The noise produced by such forces is generally reduced by the proper design of motor parts so as to separate, as much as is practicable, force frequencies and motor resonances.

Forces caused by motor dissymmetries can also act to produce directly radiated motor noise, especially in the vicinity of the rotor and stator resonant frequencies. However, more important from the standpoint of vibration coupling to the appliance, such dissymmetries often introduce unbalanced air-gap forces causing rotor flexure, especially in the vicinity of rotor resonance (usually 400 to 800 cps). Such rotor vibration can in turn flex the end bells. Thus, a mounting system utilizing the through bolts can, if not properly designed, transfer this vibratory energy to the appliance.

The torque pulsation forces generated by the rotor current and air-gap flux cause the rotor and stator to vibrate relative to each other about the shaft axis at a frequency of 2*f* cps. Since the stator is usually round, and concentric to the shaft. the torsional vibration of the motor alone acts in shear on the air and hence is not radiated as sound. Surfaces perpendicular to the direction of vibration are too small to radiate effectively the low frequency involved. However, this vibration can be effectively coupled to the appliance and hence radiated as sound if proper vibration isolation is not provided.

The torque pulsation forces acting on the rotor can have an axial component of force unless the skew is of a herringbone pattern. This axial component causes the rotor to push against the motor-thrust bearing. If the motor has free axial end play relative to the end shields, this axial force may cause the rotor to vibrate axially, transmitting vibration from the end of the motor to the supports and the appliance.

The solenoid force acting on the rotor and stator behaves in much the same manner as does the axial component of the torque pulsations, and is usually kept low in amplitude by careful rotor-to-stator axial alignment and/or by the proper design of rotor-thrust collar and end-shield thrust surfaces.

Methods for Reducing Amount of Vibration Transmission From Motors to Appliances

FUNDAMENTAL METHODS

Some methods for reducing the amount of vibration that is transmitted from a motor to the associated appliance are as follows:

1. To increase the mass of the part of the motor to which the forces are applied.
2. To add a tuned vibration absorber to the motor or vibrating part.
3. To add an inertial-friction vibration damper to the motor.
4. To use vibration isolation between the motor and the appliance.
5. To eliminate any vibration caused by too low a stiffness of the "rigid" type of motor supports.

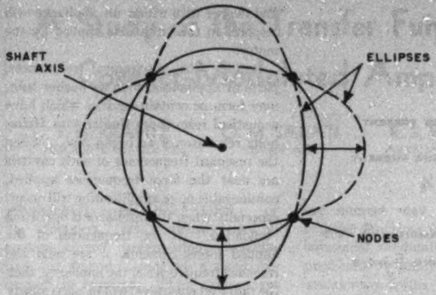


Fig. 7. Four-node mode of stator vibration in radial plane

Description of Methods and Practical Design Factors Involved

The possibility of reducing the amplitude of the applied forces should not be overlooked. Changes in the electrical design of a motor to obtain modified 2-phase operation with a run capacitor will reduce torque pulsation forces. However, the capacitor is usually more expensive and takes more space than do isolation mountings, and the maximum reduction is limited.

The possibility of increasing the mass of the motor is opposite to the basic trend in design which results in smaller and lighter motors to obtain greater output per unit of cost. Increased mass is expensive and space-consuming in relation to the improvement obtained. Vibration reduction obtained in this manner is proportional to the ratio of the masses, except for resonance effects.

The application of a tuned vibration absorber may be possible, a device which opposes any vibration of the body to which it is attached. The tuned absorber is essentially a spring and a mass which are attached to the motor so that it is free to vibrate in the same direction as the motor for the vibration which is to be reduced. The natural frequency of this system is made the same as the frequency of the applied force, but the mass vibrates 180 degrees out of phase, so that the applied force is opposed by an equal and opposite force and canceled out. The theory shows that the mass of the absorber is not important since generally the assumption is that the mass can attain an amplitude of vibration which allows its acceleration forces to equal the applied forces. However, practical springs have limiting stresses and must have a constant large enough to allow this force at stresses below the yield point. This value of spring constant dictates the size of the mass since the resonant frequency is given. The advantage of the well-designed tuned vibration absorber is that it eliminates practically all of the vibration of the body to which it is attached for a particular frequency.

The disadvantages of the tuned absorber, relative to vibration isolation, are numerous. Space and material are required with resulting higher cost; long fatigue life is not easily obtained; the device must be tuned accurately to the applied force frequency; inherent damping must be low, and resonant frequencies on both sides of the force frequency are created. Also, the absorber must be designed to withstand the peak vibration which it may have to counteract, or it may be damaged.

The inertial-frictional vibration damper is another possibility. This consists of a mass fastened to the motor in such a way that relative movement is possible but accompanied by frictional resistance which converts vibrational energy to heat. This device is in a class with the tuned absorber but is less versatile and suffers from a limited life due to possible wear unless fluid friction is used. It requires and may lose adjustment, works only above a minimum value of the applied force, and is relatively expensive.

The next method is vibration isolation, which consists of supporting the motor relative to the appliance with springs. Thus, the motor is allowed to vibrate in the direction of the applied force allowing the inertia forces of the vibrating motor, as well as the supporting springs, to oppose the applied forces. As the freedom for vibration increases, the inertia forces increase, thus decreasing the transmitted vibration. Isolation involves the transfer of vibratory energy between the motor and the springs as opposed to the absorption of energy and conversion to heat involved in damping. The springs used to provide the freedom to vibrate are usually called isolation mountings.

The degree of isolation obtained is defined in terms of the transmissibility T, which is the ratio of the amplitude of the force transmitted by the mountings to to amplitude of the force applied to the motor. The equation for T when damping in the mountings is neglected is (assuming that the point on the appliance to which the flexible mounting system is attached to be perfectly rigid)

$$T=\frac{1}{r^2-1} \tag{1}$$

in which r, the isolation ratio, is the ratio of the frequency of the disturbing force f_a, to the natural frequency of the spring mass system formed by the mountings and the motor f_n. The equation for f_n is

$$f_n=1/2\pi\sqrt{K/M} \tag{2}$$

for linear vibrations, and, for torsional vibrations

$$f_n=1/2\pi\sqrt{K_T/I} \tag{2A}$$

where K is spring constant of mounting, M is mass, and I is polar moment of inertia of the motor. A plot of the transmissibility T, for various values of the frequency ratio r, is given in Fig. 8.

Since T depends on f_n, it is a function of K and of M, the motor mass, so that the isolation provided by a mounting only has meaning in connection with a given mass. Actual mountings all have some hysteresis or damping ranging from a very low value for steel springs to relatively high value for rubber and cork. The transmissibility with damping is given by the equation (assuming that the point on the appliance to which the flexible mounting system is attached to be perfectly rigid)

$$T=\left[\frac{1+\left(\frac{2c}{c_c}r\right)^2}{(1-r^2)^2+\left(\frac{2c}{c_c}r\right)^2}\right]^{1/2} \tag{3}$$

where c = damping factor and c_c = critical damping.

This equation in effect says that, with damping, T depends on the frequency ratio r, and the damping factor. The dashed curve in Fig. 8 gives the value for T for a value of damping commonly used in rubber motor mountings and shows that for frequency ratios greater than 1.414, that damping increases transmissibility or decreases isolation; but for lower frequency ratios in the amplification range, it reduces transmissibility toward the value of 1.0 obtained by a perfectly rigid mounting. The decreased isolation caused by damping can be compensated

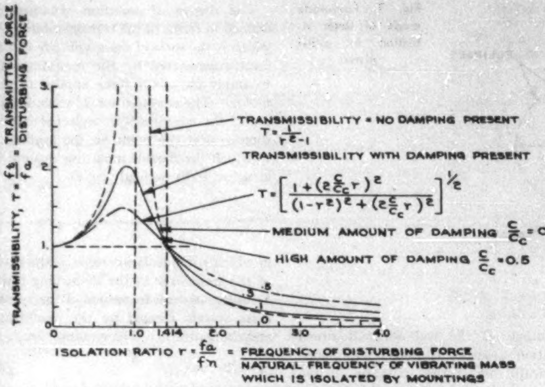


Fig. 8. Transmissibility and isolation ratio

for by a slight increase in the isolation ratio as shown by Fig. 8.

Provision of adequate isolation of all the motor forces from the appliance is obtained by providing a value of K in the mountings supporting the motor for each direction of vibration which will yield a resonant frequency sufficiently lower than the frequency of the force causing the vibration. In practice, other requirements on the mountings, such as resistance to belt pull, must be considered.

The degree and direction of isolation required for each disturbing force depends on several things: 1. the efficiency of the appliance as a radiator of received vibration, 2. the relation of the frequencies of vibration transmitted to the appliance to the resonant frequencies of the appliance, 3. the allowable radiated noise level for the appliance, and 4. the magnitude of the disturbing forces in the motor.

The effectiveness of vibration mountings can be increased by attaching the mountings to nodal points or points on the motor with the least vibration, such as close to the shaft for the torsional mode. Mountings should be attached symmetrically about or in the plane of the center of gravity of the motor to keep the possible degrees of freedom to a minimum and the mounting simple and low in cost.

The last method of reducing the magnitude of vibration transmitted from the motor to the appliance is the elimination of vibration amplification which is distinguished from vibration isolation, in that vibration isolation reduces the transmissibility below 1 (the transmissibility for a rigid mounting), whereas reduction of vibration amplification reduces the transmissibility from values appreciably greater than 1 to that of 1. This distinction is made because what appears to be a rigid mounting at first glance may actually be a mounting with sufficient compliance to provide an amplification of vibration over that of a rigid mounting, and in such cases stiffening of the mounting will provide an appreciable and possible sufficient reduction in transmitted vibration. Attainment of a rigid mounting increases the effective mass of the motor and thus reduces vibration by the increase in mass obtained. In general, however, a truly rigid mounting system is difficult to attain since, from a vibration standpoint, the appliance itself must be considered a series of spring-coupled masses.

Relation of Vibration Response Characteristics of Appliance and Associated Motor

Each part of every appliance has various mechanical resonant frequencies and these are determined by manufacturing variations as well as by design. The sheet-metal housings and tubs common to many appliances have a wide variety and number of resonant frequencies, and these vary from machine to machine. Slight variations which are otherwise insignificant can cause appreciable variations in the resonant frequencies of flat plates. The point at which vibration is applied and the stress in the part can also vary the response characteristics.

The use of vibration isolation between assemblies, damping materials, and the type of connections also affect the manner and degree with which an appliance will respond to the vibrations applied by the motor.

Besides the mechanical resonances, parts of appliances, such as washer tubs, may form or contain cavities which have acoustical resonances similar to a Helmholtz resonator or an organ pipe. When the resonant frequencies of such cavities are near the force frequencies applied, considerable noise amplification will occur, especially when the appliance is operating "unloaded." If the frequencies of the applied forces coincide or are near the resonant frequencies of the appliance, then the radiated noise level will be particularly high unless the vibrations transmitted to the appliance are very low, such as those obtained by a high-isolation ratio.

Many of these resonant frequencies are difficult to predict in the design stage. Of course, inherent stiffness or damping can be designed into various critical parts within certain limits.

The wide range and high amplitude of vibrations which can be transmitted from a motor to the appliance and the inherent sensitivity of many parts of appliances to a wide range of vibrations of relatively low magnitude results in a relation between the motor and appliance of a complex nature. Thus, the noise level of an appliance is often difficult to control and predict from unit to unit. However, if one of the available methods for reducing the amount of vibration transmitted from the motor to the appliance, particularly isolation, is effectively utilized, then the noise level radiated by the appliance due to motor vibration can be kept lower and more uniform in value. The design of the simplest and lowest-cost system which will isolate all the motor forces from the appliance, to the degree required to obtain a balance between noise caused by motor vibration and that caused by other sources, requires the appliance and motor designer to meet on a common ground. The best design of an isolation system is the one which integrates the peculiarities of both motor and appliance design with realistic requirements for both so that a smooth transition from motor to appliance is evidenced.

References

1. A Loudness Chart for Octave-Band Data on Complex Sounds, F. Mintz, F. G. Tyzzer. *Journal of the Acoustical Society of America*, New York, N. Y., vol. 24, pp. 80–82.

2. Single-Phase Induction Motor Noise Due to Dissymmetry Harmonics, D. F. Muster, G. L. Wolfert. *AIEE Transactions*, vol. 74, pt. III, 1955 (Feb. 1956 section), pp. 1365–72.

3. The Revolving Field Theory of the Capacitor Motor, Wayne J. Morrill. *Ibid.*, vol. 48, Apr. 1929, pp. 614–29.

A Study of the Transfer Function of Contact-Modulated Amplifiers

F. H. KRANTZ
ASSOCIATE MEMBER AIEE

O. M. SALATI
MEMBER AIEE

R. S. BERKOWITZ
ASSOCIATE MEMBER AIEE

CONTACT MODULATORS, or "choppers," are frequently used in modern circuit design where modulation (or demodulation) at a fixed frequency is required. Their advantages for this service include extremely high ratios of open-circuit to closed-circuit resistance, almost negligible internal noise, and excellent stability and reproducibility. These properties make the contact modulator particularly suitable for use in several types of d-c and a-c amplifiers as well as in d-c servomechanisms employing a-c control elements.

Basically, the contact modulator is a single-pole switch, driven at a fairly low audio rate. As the input element of an a-c coupled amplifier, the modulator provides a means of converting d-c and low-frequency signals to signals which lie within the pass band of the amplifier. After one or more stages of voltage amplification, the signal may be demodulated in a second converter to produce a high-level output having essentially the input wave form. Over-all d-c feedback is often employed to obtain gain stabilization.

Various adaptations of this general approach may be found in commercially available d-c microvoltmeters and micromicroammeters and in the familiar carrier-type servomechanism. In the latter application, demodulation is effected through the use of a 2-phase induction motor.

Some aspects of the over-all performance of modulated systems have been studied in recent years, chiefly with regard to servomechanisms.[1–5] In most, if not all, cases, sinusoidal modulation has been assumed. This has been sufficient, however, for the purpose of showing the general equivalence of such systems to pure d-c systems in many steady-state and transient characteristics.

In this paper, the transfer function derived for contact-modulated amplifiers includes both sinusoidal modulation and some types of impulse modulation. In addition, it will be shown that by proper adjustment of the contact modulator parameters, certain unique system transfer functions of substantial engineering interest may be obtained. These transfer functions are not available from sinusoidal modulation nor can they be predicted by analyses based upon sinusoidal wave forms.

The General Problem

Systems of the type considered in this paper have the general form shown in Fig. 1. In amplifiers designed for the measurement of low-level signals, the low-frequency input $e_i(t)$ is chopped at a 60-cycle rate, the resulting voltage amplified as much as 40 million times and then passed through a demodulator operating synchronously with, and identical to, the input modulator. High-order harmonics of the chopping frequency in the output signal are removed by the low-pass filter. Gain stabilization is obtained by feeding back to the input a fraction of the output voltage.

The transfer function under investigation is the relationship between the input and output voltages in the system of Fig. 1. Consider, therefore, an input voltage of the form

$$e_i(t)=A\epsilon^{j\omega_1 t} \tag{1}$$

For the purposes of this analysis, it will be sufficient to consider only those components in the output having the same frequency as the input signal. Other components of course exist but are highly attenuated by the low-pass filter. Hence

$$e_0(t)=C\epsilon^{j\omega_1 t} \tag{2}$$

The feedback voltage $e_f(t)$ is a fraction β of the output or

$$e_f(t)=\beta C\epsilon^{j\omega_1 t} \tag{3}$$

where β may be a function of the frequency ω_1. The voltage at point *1* of Fig. 1 is the sum of the input and feedback signals.

$$\begin{aligned} e_1(t)&=A\epsilon^{j\omega_1 t}+\beta C\epsilon^{j\omega_1 t}\\ &=(A+\beta C)\epsilon^{j\omega_1 t} \end{aligned} \tag{4}$$

The modulating wave form, of fundamental frequency ω_0, may be a simple sinusoid as assumed in some analyses, or may be quite complex, as is often the case with contact modulators. For complete generality, therefore, the modulating wave form is best expressed as

$$P(t)=\sum_{m=-\infty}^{m=\infty}P_m\epsilon^{jm\omega_0 t} \tag{5}$$

Since, in amplifiers of this type, the modulation process is one of multiplication of wave forms, the voltage at point *2* is given by

$$e_2(t)=(A+\beta C)\sum_{m=-\infty}^{m=\infty}P_m\epsilon^{j(\omega_1+m\omega_0)t} \tag{6}$$

The a-c amplifier is assumed to have a transfer function of the form

$$K(\omega)=GZ(\omega)\epsilon^{j\theta(\omega)}$$

where both the attenuation Z and the phase shift θ are functions of the signal frequency. In modulated systems, this expression must be modified by the substitution of $(\omega_1+m\omega_0)$ for ω, giving

$$\begin{aligned} K(\omega)&=G[Z(\omega_1+m\omega_0)\epsilon^{j\theta(\omega_1+m\omega_0)}]\\ &\equiv GZ_m\epsilon^{j\theta_m} \end{aligned} \tag{7}$$

where G is the amplifier frequency-invariant gain, Z_m is defined as the attenuation to the signal of frequency $(\omega_1+m\omega_0)$, and θ is the phase shift to that signal. Accordingly,

$$e_3(t)=(A+\beta C)G\sum_{m=-\infty}^{m=\infty}P_mZ_m\epsilon^{j[(\omega_1+m\omega_0)t+\theta_m]} \tag{8}$$

The output modulator may also have a complex wave form as indicated by

$$Q(t)=\sum_{n=-\infty}^{n=\infty}Q_n\epsilon^{jn\omega_0 t}$$

In addition, it may be shifted in phase with respect to the input modulator. If the phase angle has a magnitude of ϕ radians (lagging), a more complete expression for $Q(t)$ would be

$$Q(t)=\sum_{n=-\infty}^{n=\infty}Q_n\epsilon^{jn(\omega_0 t-\phi)} \tag{9}$$

The signal at point *4* is therefore

$$\begin{aligned} e_4(t)=(A+\beta C)G&\sum_{m=-\infty}^{m=\infty}\sum_{n=-\infty}^{n=\infty}\times\\ &P_mQ_nZ_m\epsilon^{j[(\omega_1+m\omega_0+n\omega_0)t+\theta_m-n\phi]} \end{aligned} \tag{10}$$

Paper 57-204, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted July 3, 1956; made available for printing January 9, 1957.

F. H. KRANTZ is with the Burroughs Corporation, Paoli, Pa., and O. M. SALATI and R. S. BERKOWITZ are with The Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

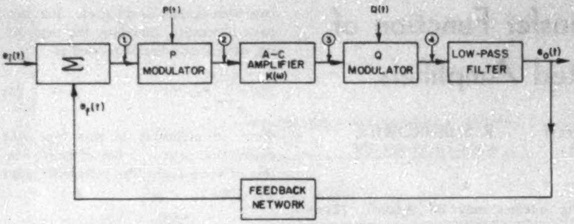


Fig. 1. Block diagram of a contact-modulated feedback amplifier

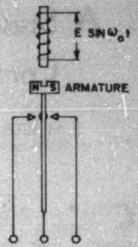


Fig. 2. Simplified structure of a contact modulator

Attention should now be directed at those components in the voltage $e_4(t)$ of frequency ω_1. Their magnitude and phase angle can be found by letting $n=-m$, which yields

$$e_4(t,\omega_1)=(A+\beta C)G\times \sum_{m=-\infty}^{m=\infty} P_mQ_{-m}Z_m\epsilon^{j(\omega_1t+\theta_m+m\phi)} \quad (11)$$

For convenience, it will be assumed that for low frequencies the output filter contributes negligible attenuation and phase shift. In those cases in which these characteristics take on substantial values, the relationships below may be modified in a straightforward manner.

With the foregoing assumption, $e_0(t)=e_4(t)$ and the two expressions for output voltage may be equated.

$$C\epsilon^{j\omega_1t}=(A+\beta C)G\times \sum_{m=-\infty}^{m=\infty} P_mQ_{-m}Z_m\epsilon^{j(\omega_1t+\theta_m+m\phi)}$$

The over-all transfer function is found to be

$$\frac{C}{A}=\frac{G\sum_{m=-\infty}^{m=\infty} P_mQ_{-m}Z_m\epsilon^{j(\theta_m+m\phi)}}{1-\beta G\sum_{m=-\infty}^{m=\infty} P_mQ_{-m}Z_m\epsilon^{j(\theta_m+m\phi)}} \quad (12)$$

It should be carefully noted that although the input frequency does not appear explicitly in equation 12, Z_m and θ_m are functions of ω_1 by equation 7.

Equation 12 can be compared directly with the standard expression for the over-all gain of a conventional feedback amplifier,

$$\text{Net gain}=\frac{\mu}{1-\mu\beta} \quad (13)$$

from which it can be seen that the contact-modulated amplifier differs only in the forward path characteristics, as expected. To the forward path, therefore, let there be assigned an equivalent μ such that

$$\mu_{\text{equiv}}=Z'\epsilon^{j\theta'}=G\times \sum_{m=-\infty}^{m=\infty} P_mQ_{-m}Z_m\epsilon^{j(\theta_m+m\phi)} \quad (14)$$

where Z' is the forward attenuation and θ' is the forward phase shift.

It is now clear that if the forward path characteristics can be determined for a given set of parameters, the closed-loop transfer function can easily be found as for a conventional amplifier. The remainder of this paper will therefore be devoted to a study of μ_{equiv}, or, in terms of Fig. 1, the transfer function between point *1* and the output.

Modulator Characteristics

Contact modulators or choppers may be represented electrically as single-pole double-throw switches operating in synchronism with the drive coil voltage as shown in Fig. 2. A typical chopper wave form is shown in Fig. 3. In the latter figure it can be seen that the wave form is essentially square except for periods of nonconduction when the armature is in transition from one contact to the other. Alternation of polarity is effected by using the modulator in conjunction with an input transformer having a center-tapped primary winding.

The average or d-c value of the wave form of Fig. 3 is easily found from

$$P_0=\frac{1}{T}\int_0^T P(t)dt = \frac{(a_3+a_4)-(a_1+a_2)}{2} \quad (15)$$

The other coefficients are determined from

$$P_m=\frac{1}{T}\int_0^T P(t)\epsilon^{-jm\omega_0t}dt = \frac{1}{2jm\pi}[\epsilon^{-jma_1\pi}-\epsilon^{jma_2\pi}+\epsilon^{jma_4\pi}-\epsilon^{-jma_3\pi}] \quad (16)$$

for m even, and

$$=\frac{1}{2jm\pi}[\epsilon^{-jma_1\pi}+\epsilon^{jma_2\pi}+\epsilon^{jma_3\pi}+\epsilon^{jma_4\pi}] \quad (17)$$

for m odd. Also

$$P_{-m}=\frac{1}{T}\int_0^T P(t)\epsilon^{jm\omega_0t}dt \quad (18)$$

showing that P_{-m} is simply the conjugate of P_m.

Two special cases occur frequently in the examples which follow. The first is the ideal full-wave modulator, for which $a_1=a_2=a_3=a_4=0$. In this case, $P_0=0$ and $P_m=2/jm\pi$ for m odd. The other is the ideal half-wave modulator, characterized by $a_1=a_2=0$; $a_3=a_4=$

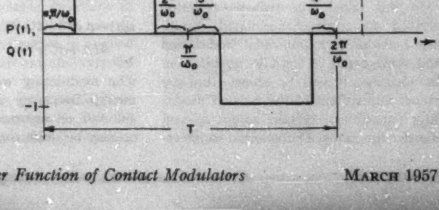


Fig. 3. Typical modulation wave form

0.5. For this case $P_0=0.5$ and $P_m=1/jm\pi$ for m odd. In both cases $P_m=0$ for m even.

Other special cases, involving various kinds of symmetry often found in practice, may be easily derived from the foregoing equations 15, 16, and 17 if required.

Simplified Transfer Functions of the Forward Path

It was shown in the previous section that the transfer function of the forward path of contact-modulated amplifiers was dependent upon the characteristics of both modulators and the amplifier. In order to demonstrate in simplest form the relationship between modulator parameters and effective gain, it will be assumed in this section that the amplifier attenuation (Z_m) is unity and the phase shift (θ_m) is zero. Then

$$Z'e^{j\theta'}=G\sum_{m=-\infty}^{m=\infty} P_m Q_{-m} e^{jm\phi} \tag{19}$$

More practical amplifiers will be considered in the following section.

The simplest of cases is that in which both modulators are ideal full-wave with $\phi = 0$. Then, from equation 17,

$$P_m=\frac{2}{jm\pi},\ Q_{-m}=\frac{2j}{m\pi}\ (m \text{ odd})$$

$$Z'e^{j\theta'}=G\cdot\frac{8}{\pi^2}\sum_{m=1}^{m=\infty \text{ odd}} 1/m^2 \tag{20}$$

From the formula for Bernouli's numbers with $n = 1$,

$$B_1=1/6=\frac{2\cdot 2}{3\cdot\pi^2}\left[1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots\right]$$

or

$$\sum_{m=1}^{m=\infty \text{ odd}} 1/m^2=\pi^2/8 \tag{21}$$

Substituting equations 20 and 21 in equation 19

$$Z'e^{j\theta'}=G \tag{22}$$

as expected.

A second case of importance is that in which the input modulator is ideal half-wave with an ideal full-wave output modulator. For this case,

$$P_0=1/2,\ P_m=\frac{1}{jm\pi}\ (m \text{ odd})$$

$$Q_0=0, Q_{-m}=\frac{2j}{m\pi}\ (m \text{ odd})$$

and

$$Z'e^{j\theta'}=G\cdot\frac{4}{\pi^2}\sum_{m=1}^{m=\infty \text{ odd}} 1/m^2=G/2 \tag{23}$$

The inverse case of a full-wave input modulator and a half-wave output modulator obviously produces the same result.

It will now be instructive to find the effective gain when both modulators are ideal half-wave. From the previous equations,

$$P_0=1/2;\ P_m=1/jm\pi (m \text{ odd});\ Q_0=1/2;\quad Q_{-m}=j/m\pi\ (m \text{ odd})$$

$$Z'e^{j\theta'}=G\left[P_0Q_0+\sum_{m=1}^{m=\infty \text{ odd}} P_m Q_{-m}\right]$$

$$=G\left[1/2\cdot 1/2+\frac{2}{\pi^2}\sum_{m=1}^{m=\infty \text{ odd}} 1/m^2\right]=G/2 \tag{24}$$

In comparing the second and third examples, it should be carefully noted that although the equivalent gain is the same in the two configurations, there is a marked difference in the way in which input signals are transmitted through the system. In the third case, half of the information is carried in side bands while the other half is transmitted via the average component of the modulator wave forms. In case 2, $P_0Q_0=0$ and the signal is transmitted through the amplifier entirely in side bands. This difference will become especially important when practical amplifiers are considered.

In addition to the foregoing equations 22, 23, and 24, other useful relationships may be found for the simplified amplifier considered in this section. These are given in the following without proof, although the derivations are straightforward.

1. If one of the modulators is ideal full-wave and the other is completely asymmetrical, the gain is

$$Z'e^{j\theta'}=G\left(1-\frac{a_1+a_2+a_3+a_4}{2}\right) \tag{25}$$

2. If both modulators are asymmetrical, the gain becomes

$$Z'e^{j\theta'}=G\left(1-\frac{A_1+A_2+A_3+A_4}{2}\right) \tag{26}$$

where A_1 is equal to the larger of the two parameters defining the dead time for the first quadrant, etc.

3. If a phase shift ϕ exists between two ideal full-wave modulators such that the output modulator lags the input, the gain is found to be

$$Z'e^{j\theta'}=G\left(1-\phi/\frac{\pi}{2}\right) \tag{27}$$

It is interesting to note that in all cases considered in this section, the equivalent gain is given by G times the average value of the wave form which is the product of the two modulation wave forms. The phase shift is, of course, zero.

Transfer Functions of Resistance-Capacitance Coupled Amplifiers

A study of contact-modulated d-c amplifiers shows that one of their principal advantages is that they permit the amplification of d-c and low-frequency signals by an amplifier having poor low-frequency characteristics. By this means, highly precise measurements can be obtained at a fraction of the cost of pure d-c systems of comparable quality. In the study of practical amplifiers, therefore, it may be expected that the frequency response will not be flat, as assumed in the previous section for purposes of illustration, but will exhibit marked loss of gain and appreciable positive phase angle in the low-frequency range. High-frequency attenuation and negative phase shifts are also present but in a frequency range which has no appreciable affect on the action of the contact modulators.

The simplest impedance function found in a-c amplifiers is that representing a single resistance-capacitance interstage coupling network. The transfer function (e_0/e_i) of such a network can be written as

$$\frac{e_0}{e_i}=\frac{1}{\sqrt{\left(\frac{\omega_c^2}{\omega}+1\right)}}\ \text{where}\ \begin{array}{l}\omega_c=\frac{1}{\tau_c}=\frac{1}{RC}\\ \theta=\tan^{-1}\frac{\omega_c}{\omega}\end{array} \tag{28}$$

Plots of attenuation and phase shift as a function of the ratio ω/ω_c are given in Fig. 4.

In modulated systems, the frequency ω must be replaced with $(\omega_1 + m\omega_0)$ by equations 6 and 7. Letting

$$\omega_1=k\omega_0\ (0\leqslant k<1/2)$$

obtains

$$\frac{\omega_c}{\omega}=\frac{\omega_c}{\omega_1+m\omega_0}=\frac{\omega_c}{\omega_0(k+m)}$$

and

$$\frac{e_0}{e_i}=\frac{e^{j\theta}}{\sqrt{\left[\frac{\omega_c}{\omega_0(k+m)}\right]^2+1}}$$

$$\text{where } \theta=\tan^{-1}\frac{\omega_c}{\omega_0(k+m)} \tag{29}$$

If the expression for the a-c amplifier transfer function given by equation 29 is substituted for $Z_m e^{j\theta_m}$ of equation 14, it will be possible to find the over-all transfer function of the forward path as a func-

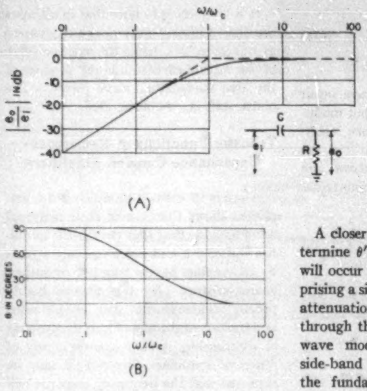


Fig. 4. Transfer function of a resistance-capacitance coupling network

A—Amplitude characteristic
B—Phase characteristic

tion of any of the system variables. Of particular interest is the variation in the forward path attenuation and phase shift as a function of signal frequency and amplifier time constant. These functions are shown in Fig. 5 for the case of ideal full-wave input and output modulators having zero relative phase shift.

In Fig. 5, the magnitude and phase angle of the equivalent gain are plotted for values of signal frequency between 0 and $\omega_0/2$ with the ratio ω_c/ω_0 taken at selected values in the range 0.1–2. From this figure, several characteristics of the forward transfer function are immediately apparent.

1. For a given amplifier time constant, the gain is substantially independent of signal frequency.
2. The phase shift to d-c signals is unconditionally zero.
3. At input frequencies not equal to zero, the over-all phase shift θ' may be entirely positive, entirely negative, or positive in some frequency regions and negative in others.

A closer study of the factors which determine θ' shows that a net phase shift will occur whenever the two signals comprising a side-band pair are given differing attenuations and phase shifts in passing through the a-c amplifier. With square-wave modulators the most important side-band pair is of course that around the fundamental modulator frequency. It is also possible to show that, to a very good approximation, the phase shift to low-frequency signals will be negative if $\omega_c < \omega_0$ and positive if $\omega_c > \omega_0$. This observation is borne out by the curves of Fig. 5.

In the design of contact-modulated d-c amplifiers, it is customary to introduce a phase difference between modulators in such a direction as to compensate for the phase shift produced in the a-c amplifier. The amount of phase difference required can be found by substituting various values of ϕ into equation 14.

Assuming ideal full-wave modulators for convenience, the effect of varying ϕ between 10 and -50 degrees is shown in Fig. 6 for $\omega_c = 2\omega_0$. The compensating effect of shifting the phase of one modulator with respect to the other is clearly indicated. It should be noted, however, that a comparatively large phase difference (approximately 40 degrees in this case) is required to provide reasonably complete compensation. Also of interest is the rather large increase in θ' for a small phase-angle difference between modulators in the opposite direction.

Throughout the previous discussion, ideal full-wave modulators have been assumed so that the effects of other system variables might be clearly shown. In practice, however, no real modulator meets this specification since dead-time and asymmetry are always present, either initially or after a brief period of operation. In most cases, small departures from the ideal conditions produce proportionately small changes in Z' and θ'. The limiting case of dead-time and asymmetry, i.e., the half-wave modulator, deserves special attention, however.

The half-wave modulator differs significantly from the full-wave modulator in two important respects. First, its gain is only 50 per cent of that of the full-wave modulator. Second, the Fourier expansion of this wave form has the maximum possible average value, 1/2. It is the effect of the large average value with which we are concerned.

In carrier-type amplifiers, an input modulator having asymmetry about π/ω_0 will permit a fraction of the input signal to pass through the input converter unmodulated. This fraction is the average value of the modulating wave form. Since the input signal is typically of low frequency, it will be considerably attenuated and shifted in phase by the poor low-frequency response of the a-c amplifier. If now the output modulator is completely symmetrical, the low-frequency signal at the amplifier output will be completely converted to side-band frequencies lying

Vol.
Pt. 2
1957

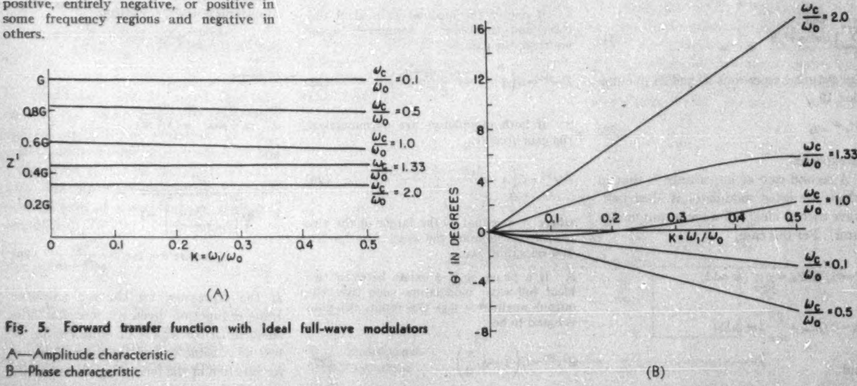
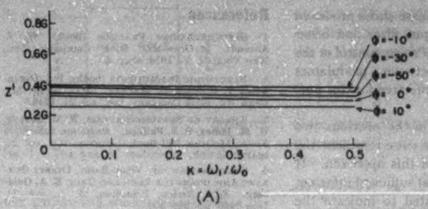


Fig. 5. Forward transfer function with ideal full-wave modulators

A—Amplitude characteristic
B—Phase characteristic



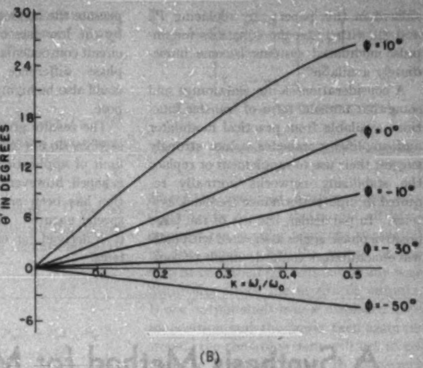


Fig. 6. Forward transfer function with a phase angle (ϕ) between modulators ($\omega_c = 2\omega_0$)

A—Amplitude characteristic
B—Phase characteristic

outside of the pass band of the low-pass output filter. In this case, the effect of asymmetry in the input modulator is simply to produce a reduction in Z'. On the other hand, if the output modulator is not symmetrical, some fraction of the amplifier low-frequency output will pass through the output modulator and low-pass filter to combine with the demodulated side bands considered previously. Since the gain to the latter signal is reduced by a factor of four in the worst case because of asymmetry, the effect of adding the "direct" component to it may be quite significant.

If both the input and output converters are ideal half-wave with zero relative phase difference, the effective forward gain is then given by the following equation:

$$Z'e^{j\theta'} = \frac{G}{4}\left[Z_{\omega 1}e^{j\theta\omega_1} + \frac{4}{\pi^2}\sum_{m=1}^{m=\infty\,\text{odd}} \times 1/m^2\left(Z_{m\epsilon}e^{j\theta_m} + Z_{-m\epsilon}e^{j\theta_{-m}}\right)\right] \quad (30)$$

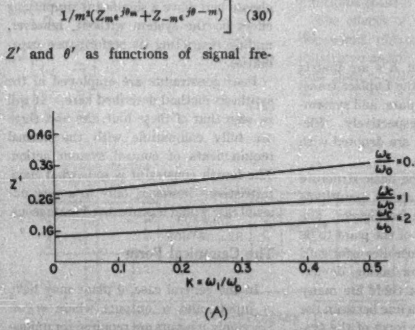
Z' and θ' as functions of signal frequency are shown graphically in Fig. 7 for $\omega_c = 2\omega_0$, ω_0, and $1/2\omega_0$. The effect of the "direct" component upon θ' is quite evident, producing in each case a very substantial positive addition to the phase angle obtained for ideal full-wave modulators.

Summary and Conclusions

It has been the purpose of this paper to examine briefly the transfer function for contact-modulated systems. In general, these relationships have not heretofore been widely available, even for the simplified case of sinusoidal modulators. It is hoped that the more exact functions presented may serve as a guide to the circuit designer to obtain higher system performance.

One of the results of the foregoing work has been to demonstrate the usefulness of the complex Fourier series in the analysis of systems employing modulation by nonsinusoidal wave forms. While not the most sophisticated of available techniques, this approach offers the advantages of mathematical simplicity, ease of interpretation, and the fact that it is widely employed.

The closed-loop transfer function derived under "The General Problem" provides a means for determining Nyquist stability, transient response, over-all gain and phase shift, and other performance characteristics for modulated feedback systems. Although somewhat similar results have recently appeared in the literature for some types of sampled-data systems, such studies have been solely concerned with a single impulse-function modulator, permitting considerable simplification in mathematical operations. In this connection, it should be observed that many sampled-data systems are in actuality a special case of the system con-



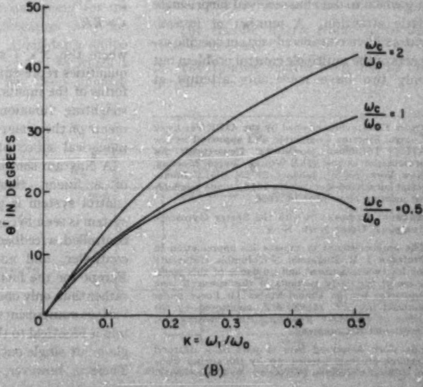


Fig. 7. Forward transfer function with half-wave modulators

A—Amplitude characteristic
B—Phase characteristic

sidered in this paper; by replacing P_m and Q_n with $1/2\pi$, the equations for impulse-modulated systems become immediately available.

A consideration of the wide range and somewhat unusual form of transfer functions available from practical modulator and amplifier parameter values strongly suggest their use to supplement or replace the stabilizing networks normally required in high-performance feedback systems. In particular, the use of the large positive phase angles associated with half-wave modulators could do much to compensate the lagging phase shifts produced by the low-pass output filter and other circuit components. Proper control of the phase difference between modulators could also be highly effective for this purpose.

The results given in the previous two sections do not in any sense represent the limit of application of this approach. It is hoped, however, that sufficient information has been presented to indicate the special nature of the transfer function as it is derived for contact-modulated systems.

References

1. SERVOMECHANISM PRACTICE (book), W. R. Ahrendt. McGraw-Hill Book Company, Inc., New York, N. Y., 1954, chap. 5.

2. ELECTRONIC INSTRUMENTS (book), I. A. Greenwood, V. Holdam, D. MacRae. *MIT Series*, McGraw-Hill Book Company, Inc., vol. 21, 1948.

3. THEORY OF SERVOMECHANISMS, N. B. Nichols, H. M. James, R. S. Phillips. *Radiation Laboratory Series*, McGraw-Hill Book Company, Inc., vol. 25, 1947.

4. STABILIZATION OF WIDE-BAND DIRECT-CURRENT AMPLIFIERS FOR ZERO AND GAIN, E. A. Goldberg. *RCA Review*, Princeton, N. J., vol. 11, June 1950, pp. 296–300.

5. D-C AMPLIFIER STABILIZED FOR ZERO AND GAIN, A. J. Williams, Jr., R. E. Tarpley, W. R. Clark. *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 47–57.

A Synthesis Method for Multipole Control Systems

HERBERT FREEMAN
MEMBER AIEE

Synopsis: A method is developed for the systematic synthesis of multipole control systems which is based on the application of certain system constraints. These constraints reduce the degrees of freedom of the system without, however, resulting in an undue restriction on system performance. A system transfer matrix is described which characterizes the system. Through the application of the constraints it becomes possible to solve this matrix for the controller transfer functions required to achieve specified performance. This permits the synthesis of multipole control systems in a manner similar to that presently possible for single-input single-output system.

THE FIELD of multipole control systems is one area of control engineering which to date has received surprisingly little attention. A number of investigators have considered certain specific aspects of the multipole control problem but only two have made any attempt at developing an over-all synthesis procedure.[1–5]

A multipole control system is a multipole[6] designed for the purpose of achieving precise, stipulated control action through the use of internal feedback. For the general multipole control system of Fig. 1

$$c_j=f_j(r_1, r_2, \ldots, r_m),\ j=1, 2, \ldots, n \tag{1}$$

If consideration is limited to time-invariant, linear input-output relations, it is possible to characterize the system by a transfer matrix. Thus

$$C_j=\sum_{i=1}^{m} K_{ji}R_i \tag{2}$$

or

$$C=KR \tag{3}$$

where the R's, C's, and K's are matrix quantities representing the Laplace transforms of the inputs, outputs, and system-weighting functions, respectively. Elements of these matrices are denoted with numerical subscripts.

A diagram showing the basic structure of a linear, time-invariant multipole control system is given in Fig. 2. The system is seen to consist of the plant to be controlled, a feedback controller, a forward controller, and an error-taking device. Except for the fact that there are many rather than only one data line between the various components, the form of the system is identical to that of the familiar single-input single-output servomechanisms. There is, however, one important difference. Since the multipole plant effects a vector transformation between its inputs and outputs, there may be as many as three distinct co-ordinate reference frames in the system. The co-ordinate frame of the system inputs may differ from that of the plant inputs, and both may differ from that of the plant outputs. The special case where D, P, and H are all diagonal matrices is, of course, trivial.

In the synthesis problem, a multipole plant is given (and assumed unalterable) and the various controllers are to be designed toward achieving a specified over-all response. For m inputs and n outputs, $m \geqq n$, there are mn distinct responses which may be specified.

System Constraints

The number of possible variables encountered in the synthesis of multipole control systems is so large that no formal, systematic synthesis procedure is possible without the imposition of some constraints on the system. The constraints chosen here have a significant simplifying effect on the system without, however, unduly restricting its performance capabilities.

Four constraints are employed in the synthesis method described here. It will be seen that of these four, the first three are fully compatible with the normal requirements of control system action. The fourth constraint is somewhat more restrictive; however, only in a very unusual case would it prove disadvantageous.

The Canonical Form

In the general case, a plant may have m inputs and n outputs where $m \geqq n$. Since only n inputs are required for unique control of the n outputs, there are m-n inputs which may be set arbitrarily. In practice, these often represent special sys-

---

Paper 57-44, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 22, 1956; made available for printing November 16, 1956.

HERBERT FREEMAN is with the Sperry Gyroscope Company, Great Neck, N. Y.

The author wishes to express his appreciation to Professor J. R. Ragazzini of Columbia University for his encouragement and guidance of this work. Some of the early portions of this research were supported by the United States Air Force under contract no. *AF 18(600)-677*, monitored by the Office of Scientific Research, Air Research and Development Command.

The work described here is part of a doctoral research program carried out by the author while on leave at Columbia University, New York, N. Y.

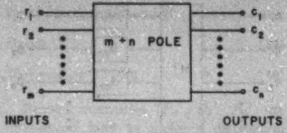


Fig. 1. The general multipole control system

tem parameters chosen to provide a desirable operating region. A generalized block diagram of a multipole control system is given in Fig. 3. By redefining the plant matrix it becomes possible to carry the *m-n* secondary inputs through the plant and obtain what might be called *m-n* "virtual" outputs in addition to the *n* real outputs. The block diagram of Fig. 4 gives the canonical form of multipole control systems upon which the synthesis method developed here is based. The canonical form of block diagram may be considered the first constraint imposed.

In Fig. 4 the data column matrices *R, E, X, Y,* and *C* are partitioned to bring into evidence the distinction between the *n* primary control lines and the *m-n* secondary lines. Thus

$$R=\begin{bmatrix} R_a \\ R_b \end{bmatrix},\quad E=\begin{bmatrix} E_a \\ E_b \end{bmatrix}, \text{ etc.} \tag{4}$$

$$C=GDE \tag{5}$$

$$E=R-Y \tag{6}$$

$$Y=HC \tag{7}$$

Solving for C,

$$C=[1+GDH]^{-1}GDR \tag{8}$$

To facilitate the solution for C_a and C_b, it is convenient to partition each of the $m\times m$ matrices D, G, and H into four submatrices, as follows:

$$D=\begin{bmatrix} D_a & D_c \\ D_d & D_b \end{bmatrix} \tag{9}$$

$$G=\begin{bmatrix} G_a & G_c \\ G_d & G_b \end{bmatrix} \tag{10}$$

$$H=\begin{bmatrix} H_a & H_c \\ H_d & H_b \end{bmatrix} \tag{11}$$

where D_a, G_a, and H_a are all $n\times n$ matrices. It follows from the definition of G and the $m\times n$ structure of P that

$$G=\begin{bmatrix} P_a & P_c \\ 0 & 1 \end{bmatrix} \tag{12}$$

Since coupling between C_a and C_b in H is meaningless from physical reasoning, one may let

$$H_c=H_d=0 \tag{13}$$

This is the second constraint. It has no significant restrictive effect on the system designer. Substituting equations 9–13 in equation 8 gives

$$\begin{bmatrix} C_a \\ C_b \end{bmatrix}= \begin{bmatrix} 1+(P_aD_a+P_cD_d)H_a & (P_aD_c+P_cD_b)H_b \\ D_bH_a & 1+D_bH_b \end{bmatrix}^{-1}\times \begin{bmatrix} (P_aD_a+P_cD_d) & P_aD_c+P_cD_b \\ D_d & D_b \end{bmatrix} \begin{bmatrix} R_a \\ R_b \end{bmatrix} \tag{14}$$

Careful interpretation of this equation is necessary because the elements involved are submatrices. Solving the equation in this form, however, yields results which are too unwieldy to be of practical use.

Independent Output Restoration

The third constraint to be imposed on the system may be described as the requirement for independent output restoration. This requirement can be stated mathematically as the condition that the solution of the homogeneous equations for the outputs, with an initial condition on one output, be zero, except for the output having the initial condition. It will be seen that this constraint, important as it may be from a mathematical point of view, imposes no significant limitation on the achievement of the objectives of a multipole control system.

The effects of the requirement for independent output restoration may be illustrated as follows. Consider a multipole shaft-positioning system to which a set of input signals has been applied. After the required settling time has elapsed, the output shafts will all have assumed the angular positions stipulated by the inputs. Let the power to the system be turned off, without in any way disturbing either the inputs or outputs. If one output shaft is now displaced from its position and the power then again restored, the control system will act to return the disturbed output to its correct position. According to the requirement for independent output restoration, the other output shafts—which were already in the correct positions—are to be completely unaffected by this.

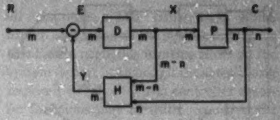


Fig. 3. General system block diagram

This limitation of output interaction should not be confused with the much more restrictive input noninteraction described by Boksenbom and Hood.[1] No limitation is imposed here on the interaction of the inputs. By means of H, any kind of linear co-ordinate transformation may be realized from the input to the output of the system.

The constraint of independent output restoration results in a significant simplification of the system equations. Consider the open-loop diagram of Fig. 5. The feedback loop of the system is shown opened at the output. According to the constraint, if a disturbance is applied to one of the set of terminals C_{ai}, an output correction force is to appear at and only at the corresponding terminal of the set of terminals C_{ao}. The same, of course, should hold for C_{bi} and C_{bo}. Hence

1. The transfer matrix from C_{ai} to C_{ao} must be diagonal.

2. The transfer matrix from C_{bi} to C_{ao} must be zero.

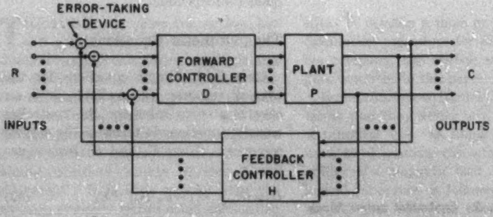


Fig. 2. Basic structure of a multipole control system

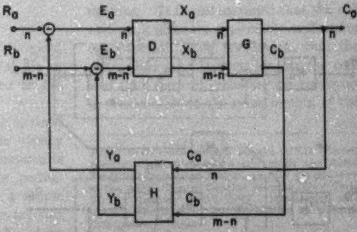


Fig. 4. Canonical form of multipole control system

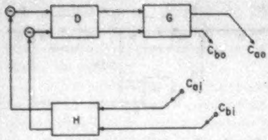


Fig. 5. System open-loop diagram

3. The transfer matrix from C_{bi} to C_{bo} must be diagonal.
4. The transfer matrix from C_{ai} to C_{bo} must be zero.

From equations 9–13 the open-loop transfer matrix is given by

$$GDH = \begin{bmatrix} (P_aD_a+P_cD_d)H_a & (P_aD_c+P_cD_b)H_b \\ D_dH_a & D_bH_b \end{bmatrix} \tag{15}$$

Hence

$$(P_aD_a+P_cD_d)H_a = \text{diagonal} \tag{16}$$

$$(P_aD_c+P_cD_b)H_b = 0 \tag{17}$$

$$D_bH_b = \text{diagonal} \tag{18}$$

$$D_dH_a = 0 \tag{19}$$

Since H_a cannot be zero,

$$D_d = 0 \tag{20}$$

Since H_b cannot be zero, except when P_c is also zero,

$$P_aD_c+P_cD_b = 0 \tag{21}$$

From equations 16 and 20,

$$P_aD_aH_a = \text{diagonal} \tag{22}$$

Equation 20 is completely under the designer's control and can be met exactly. Making use of equation 20 in the solution of equation 14 gives the following two equations for the over-all response of the system:

$$C_a = [1+P_aD_aH_a]^{-1}P_aD_aR_a + [1+P_aD_aH_a]^{-1}[P_aD_c+P_cD_b]H_b \times [1+D_bH_b]^{-1}H_b^{-1}R_b \tag{23}$$

$$C_b = [1+D_bH_b]^{-1}D_bR_b \tag{24}$$

This represents an important simplification. Equations 23 and 24 are exact and do not involve any approximation.

The System Transfer Matrix

The over-all system transfer matrix relates the outputs of the system to the inputs. Designating it by K and partitioning it in the manner described before

$$K = \begin{bmatrix} K_a & K_c \\ K_d & K_b \end{bmatrix} \tag{25}$$

and

$$C_a = K_aR_a + K_cR_b \tag{26}$$

$$C_b = K_dR_a + K_bR_b \tag{27}$$

Substituting from equations 23 and 24 in 25, one obtains

$$K = \begin{bmatrix} [1+P_aD_aH_a]^{-1}P_aD_a & [1+P_aD_aH_a]^{-1}[P_aD_c+P_cD_b] \times H_b[1+D_bH_b]^{-1}H_b^{-1} \\ 0 & [1+D_bH_b]^{-1}D_b \end{bmatrix} \tag{28}$$

It is convenient to define an additional matrix

$$M = \begin{bmatrix} M_a & M_c \\ M_d & M_b \end{bmatrix} \tag{29}$$

such that

$$M = \begin{bmatrix} [1+P_aD_aH_a]^{-1}P_aD_aH_a & [1+P_aD_aH_a]^{-1}[P_aD_c+P_cD_b] \times H_b[1+D_bH_b]^{-1} \\ 0 & [1+D_bH_b]^{-1}D_bH_b \end{bmatrix} \tag{30}$$

Noting that

$$M_c = [1-M_a][P_aD_c+P_cD_b]H_b[1-M_b] \tag{31}$$

one obtains

$$K = \begin{bmatrix} M_aH_a^{-1} & M_cH_b^{-1} \\ 0 & M_bH_b^{-1} \end{bmatrix} \tag{32}$$

$$K = MH^{-1} \tag{33}$$

By virtue of equation 32 it becomes possible to construct an equivalent system block diagram which brings into evidence the significance of the matrices M_a, M_b, and M_c, as shown in Fig. 6.

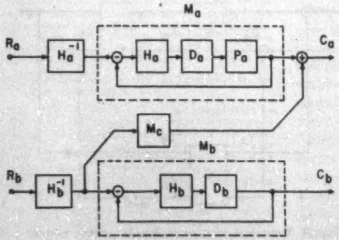


Fig. 6. Equivalent system block diagram

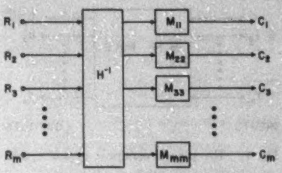


Fig. 7. Illustration of the physical significance of the H and M matrices

If equations 18 and 22 are satisfied, both M_a and M_b will be diagonal matrices. Also, if equation 21 is substituted in equation 31, M_c will vanish and the over-all system equations will reduce to

$$C_a \approx [1+P_aD_aH_a]^{-1}P_aD_aR_a \tag{34}$$

$$C_b = [1+D_bH_b]^{-1}D_bR_b \tag{35}$$

For $m=n$, equation 34 becomes exact and equation 35 vanishes.

Since M_a and M_b are diagonal and M_c is vanishingly small, M is also diagonal.

According to equation 33 then, the over-all transfer matrix K can be expressed as the product of the inverse feedback matrix H^{-1} and a diagonal matrix. This is shown graphically in Fig. 7.

Because of the complete generality of equation 33, it may be stipulated that H be a zero-memory (frequency-independent) matrix whose purpose is solely to provide the necessary input-output magnification and co-ordinate rotation. The dynamics of the system will then all be contained in the diagonal matrix M. This stipulation represents the fourth and final constraint to be imposed on the system by this synthesis procedure. It forces all input components to be subject to the same dynamics for any one output; i.e., in

$$C_j = M_{jj}\sum_{k=1}^{n}\frac{|H_{kj}|}{|H|}R_k \tag{36}$$

the dynamics of C_j for all j are determined entirely by M_{jj} since the H_{kj} are frequency-independent.

The Synthesis Equations

Once fairly simple equations for the over-all response are available, it is an easy matter to solve for the controller transfer functions in terms of the desired response. Using K_a and K_b from equation 28,

$$D_a = [P_a']^{-1}[1-K_a'H_a]^{-1}K_a' \tag{37}$$

$$D_b = [1-K_b'H_b]^{-1}K_b' \tag{38}$$

From equations 20 and 21

$$D_d=0 \tag{39}$$

$$D_c=-[P_a']^{-1}P_c'D_b \tag{40}$$

Primed quantities (P_a', P_b', P_c') are introduced here to make a distinction between the plant values used in the design (i.e., the supposed values) and the true physical values P which are never known exactly. The difference between P' and P arises from errors (or limited precision) in measurement, nonlinearities, slight time variations in P, and other approximations known or unknown. For all practical purposes, P' will be sufficiently close to P to be considered equivalent. However, in connection with stability considerations, a lack of awareness of this distinction might cause the accidental cancellation of unstable-region poles and give a false stability picture. The prime notation is, therefore, employed to indicate plant values used for the design of the controller D. Similarly, the primes on K_a' and K_b' indicate that they represent the desired rather than the actual responses.

Equations 37 through 40 uniquely determine the transfer matrix D, though without any assurance of physical realizability or system stability, which must be determined separately. It is important to note that only equation 39 can be met precisely. All others involve the plant transfer functions and can be satisfied only to the extent to which the plant transfer functions are known and unvarying, i.e., they depend on the degree of accuracy with which P' represents P.

In the important case where a one-to-one correspondence is to exist between the inputs and outputs, K' is diagonal and equations 37 and 38 reduce to:

$$(D_a)_{ij}=\frac{|P_{ji}'|}{|P_a'|}\,\frac{K_{jj}'}{1-H_{jj}K_{jj}'} \qquad i,j=1,2,\ldots,n \tag{41}$$

$$(D_b)_{ij}=\frac{K_{jj}'}{1-H_{jj}K_{jj}'} \quad \text{for } i=j \tag{42}$$

$$=0 \quad \text{for } i\neq j \qquad i,j=n+1,n+2,\ldots,m$$

Since H has been restricted to contain only zero-memory elements, its design becomes trivially simple. It is given by

$$H=[K_{ss}']^{-1} \tag{43}$$

where K_{ss}' is the specified over-all steady-state response for a unit-step input. Since K' relates the m-controlled outputs (the n physical outputs plus the m-n "virtual" ones) to the m-independent inputs, it is nonsingular and equation 43 is always valid. If K' is specified to be diagonal, H will have to be diagonal.

One might assume that an alternative approach is possible where D is fixed and H acts as a dependent design variable. It has been shown,[7] however, that this approach generally does not lead to a physically realizable system.

Conclusions

A method has been outlined which permits the solution for the controller transfer functions of a multipole control system directly in terms of the given plant transfer functions and the desired over-all responses. The constraints which had to be applied to the system to achieve this have a negligible limiting effect on the system design. Although the actual work required to solve for the controller transfer functions may be quite extensive, the procedure is straightforward and does not involve any trial-and-error approaches.

Although no mention is made here of the conditions for physical realizability and stability, work on these problems has been carried out and the results are expected to be presented later.

References

1. General Algebraic Method Applied to Control Analysis of Complex Engine Types, A. S. Bokenbom, R. Hood. *NACA Technical Report 980*, National Advisory Committee for Aeronautics, Washington, D. C., Apr. 1949.

2. Engineering Cybernetics (book), H. S. Tsien. McGraw-Hill Book Company, Inc., New York, N. Y., 1954.

3. Analysis for Control Application of Dynamic Characteristics of Turbojet Engine with Tail-Pipe Burning, M. S. Feder, R. Hood. *NACA Technical Note 2183*, National Advisory Committee for Aeronautics, Sept. 1950.

4. A Theory of Multidimensional Servo Systems, M. Golomb, E. Usdin. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 253, no. 1, Jan. 1952, pp. 29–57.

5. A Method for the Preliminary Synthesis of a Complex Multiloop Control System, D. J. Povejsil, A. M. Fuchs. *AIEE Transactions*, vol. 74, pt. II, July 1955, pp. 129–34.

6. Nonlinear Multipoles, L. A. Zadeh. *Proceedings*, National Academy of Sciences, Washington, D. C., vol. 39, no. 4, Apr. 1953, pp. 274–80.

7. The Synthesis of Multipole Control Systems, H. Freeman. *Doctoral Dissertation. Technical Report T-15/B*, Electronics Research Laboratory, Columbia University Engineering Center, New York, N. Y., Apr. 1956.

8. Multipole Sampled-Data Control Systems, H. Freeman. *Technical Report T-12/B*. *Ibid.*, Sept. 1955.

Design and Calculation of Induction-Heating Coils

R. M. BAKER
MEMBER AIEE

THROUGH the years the author has had considerable experience with induction-heating calculating methods and has studied the reaction of others confronted with the problem of designing coils and predicting performance. It is a difficult job at best when one considers the great variety of materials heated and the almost limitless variety of shapes encountered. It is little wonder that the average engineer resists using calculating methods as much as he does. In an effort to develop a more or less standard calculating procedure for internal use in his company, the author arrived at the plan described in this paper.

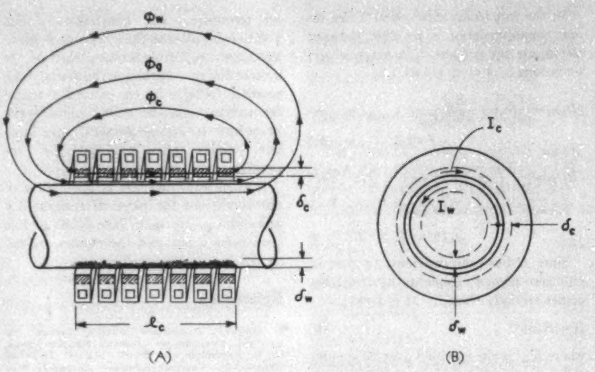
Equations are written in such general terms that they apply to a wide variety of practical jobs. A simple expression is used to determine the effective permeability of a magnetic load, and a simple logical procedure is followed to account for the end effect of electrically short coils. A coil is electrically short when an appreciable portion of its ampere-turns is used to overcome the reluctance of the external flux path and therefore is not effective in heating the load.

The method is applicable to loads of any shape so long as the coil surrounds the work piece and its length is at least equal to the minimum dimension of the coil opening. It is also assumed that the coil and work piece are reasonably uniform in cross section along the length, that the work piece is at least as long as the coil, and that the coil cross section conforms approximately to the cross section of the

---

Paper 57-4, recommended by the AIEE Electric Heating Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted June 8, 1956; made available for printing October 18, 1956.

R. M. Baker is with the Westinghouse Electric Corporation, Baltimore, Md.



(A) (B)

Fig. 1. A typical induction-heating load

work piece. The latter represents good induction-heating practice because both power factor and efficiency improve with close coupling.

The P and Q factors appearing in this paper are derived from the standard skin effect equations which have existed in the literature for the last 50 years and the author claims no originality for these. It happens that he did derive the P and Q curves for the hollow cylinder but only because these were not readily found in the literature.

In this paper, emphasis is placed on obtaining a universal method so that the user applies the same equations and same procedure in calculating a large number and variety of jobs. This will develop familiarity and facility on the part of the user, and certain trends and differences between different jobs will be immediately obvious.

In achieving the advantages of a universal method, one may sacrifice some accuracy, especially in the case of very short coils or loads of unusual shape. It is felt, however, that any accuracy sacrificed is more than offset by advantages.

The procedure in deriving equations is straightforward and should be easy to follow. Two specific calculating procedures are outlined but it should be realized that these are merely instructive and that the equations of the paper may be used in a great many other ways.

It will be assumed that the reader is reasonably familiar with electromagnetic theory and especially with skin effect phenomena and the concept of depth of current penetration. Complex phasor notation will be used generally to describe sinusoidally varying electric and magnetic quantities, complex impedances, and complex volt-amperes. A dot over a symbol indicates that the quantity generally has a real and j component. The symbol without a dot will be used to indicate the absolute magnitude of the quantity or to indicate that the quantity is in phase with the reference phasor.

The practical system of electrical and magnetic units will be used with lengths expressed in centimeters.

Nomenclature

ϕ_g = peak air-gap flux, maxwells
ϕ_w = peak flux in work piece, maxwells
ϕ_c = peak copper flux, maxwells
ϕ_0 = total peak flux linking coil turns, maxwells
A_g = cross-section air-gap area, square centimeters
A_w = cross-section area of work based on outside dimensions, square centimeters
l_c = length of coil, centimeters
H_0 = peak magnetizing force, oersteds
μ = effective magnetic permeability
P = dimensionless flux factor
Q = dimensionless flux factor
B_s = saturation flux density, gauss
δ = effective depth of current penetration, centimeters
ρ = electrical resistivity, ohm-centimeters
f = frequency, cycles per second
N_c = number of coil turns
I_c = rms coil current, amperes
k_r = empirical coil resistance factor, dimensionless
p_c = inside perimeter of coil, centimeters
E_c = rms coil voltage, volts
R_w = reflected work resistance, ohms
R_c = coil resistance, ohms
X_g = reactance due to air-gap flux, ohms
X_w = reactance due to work flux, ohms
X_c = reactance due to copper flux, ohms
Z_c = total effective coil impedance, ohms
I_1 = component of coil current in phase with H_0, amperes
I_2 = component of coil current in phase with ϕ_0, amperes
R_e = reluctance of external flux path, centimeter-gram-second units
P_w = total power input to work piece, watts
P_c = coil copper loss, watts
P_0 = coil input, watts
X_1 = (see equation 27)
X_e = end-effect reactance, ohms
R_w' = reflected load resistance corrected for end effect, ohms
X_0 = total coil reactance, ohms
k_e = end-effect reactance factor
t = wall thickness of hollow cylinder or thickness or slab, centimeters
d = cylindrical load diameter, centimeters
$(VA)_c$ = coil volt-amperes, rms volts times rms amperes
E_{c1} = rms voltage of single-turn coil or voltage per turn in multiturn winding, volts

Typical Induction-Heating Load

The load of Fig. 1 will serve as an example of the type of load considered. The cylindrical shape is shown only because it is easy to draw. The coil opening and cross section of work may have any shape as long as these are both reasonably constant along the length and provided the coil is at least as long as the minimum dimension of the coil opening.

It is convenient to start with the assumption that alternating current in the coil turns causes a field intensity, H_0, to be established in the air gap parallel to the surface of the work piece. It is also assumed that the field intensity is in phase with the reference phasor. See Fig. 3.

The total flux in the air gap is equal to the product of field intensity, H_0, and air-gap area, A_g.

$$\phi_g = H_0 A_g \tag{1}$$

In penetrating the work piece the field intensity decreases exponentially and shifts continuously in phase behind the field intensity, H_0 at the surface. The total integrated flux in the work piece consequently lags H_0 by some angle between zero and 90 degrees. This flux can be represented quite generally by the expression

$$\phi_w = \mu A_w H_0 (P - jQ) \tag{2}$$

where P and Q are dimensionless flux factors depending on load geometry, resistivity, permeability, and frequency. This equation actually defines the P and Q functions which will be discussed in more detail later. For the time being they are to be thought of simply as factors in the general expression for total work flux.

When the load is nonmagnetic, the permeability appearing in equation 2 and elsewhere in this paper is unity ($\mu = 1$). If the load is magnetic the effective permeability is dependent on H_0, and is to be calculated from the expression,

$$\mu = \frac{1.8B_s}{H_o} + 1 \qquad (3)$$

where B_s is the saturation flux density for the load material. Substituting $B_s = 18{,}000$, this becomes, for steel,

$$\mu = \frac{32{,}400}{H_o} + 1 \qquad (4)$$

This is obviously an approximation but its use is justified by experience as shown in reference 1 (see the appendix). It is adequate when H_o is equal to or greater than 60 oersteds, a condition more than satisfied in the vast majority of induction-heating applications.

The magnetic field established in the air gap also penetrates the coil turns to a certain depth and the total integrated flux in the copper can be calculated approximately from the expression

$$\phi_c = \frac{k_r p_c \delta_c}{2} H_o(1-j) \qquad (5)$$

where δ_c is the effective depth of current penetration into a flat copper surface (infinite radius of curvature), and k_r is an empirical constant introduced to account for the spacing between turns and other imperfections encountered in practical induction-heating coils.

The sum of the three component fluxes, equations 1, 2, and 5, gives the total flux ϕ_0 linking the coil turns. The magnitude and phase of this total flux in relationship to the coil current necessary to establish the field intensity H_o in the air gap serves as the basis for the present analysis.

It is customary to explain the induction-heating process by saying that the work flux induces a voltage and causes a current I_w to flow in the work as shown in Fig. 1(B). Skin effect causes this induced current to flow in a restricted manner more or less close to the surface and the resulting I^2R loss produces the desired heating. Skin effect likewise causes the coil current I_c to hug the inside of the coil turns. In so doing it encounters a resistance somewhat greater than the d-c resistance of the coil. Because the air-gap flux represents a considerable part of the total and results in purely reactive volt-amperes, the power factor of an induction-heating coil is always low, ranging from less than 0.10 to approximately 0.50 depending on frequency, closeness of coupling, and physical properties of the work piece. This simple picture of the induction-heating process is quite satisfying because it permits an understanding of what happens. This is especially true when one elaborates on the concept of depth of current penetration and how it changes with frequency and physical properties. From an analytical standpoint it is convenient, however, to base the analysis on a more general approach. The concept of depth of current penetration will be introduced only as a convenient reference dimension to be used in plotting data or writing equations.

The effective depth of current penetration into a flat surface is defined by the equation

$$\delta = \sqrt{\frac{\rho}{4\pi^2\mu f 10^{-9}}} = 5{,}040\sqrt{\frac{\rho}{\mu f}} \text{ centimeters} \qquad (6)$$

The physical constants for the coil copper may be substituted in this equation to calculate δ_c used in equation 5 or the physical constants of the load may be substituted to calculate δ_w used later in the plotting of P and Q curves.

Development of Equations

Long Coils

The total magnetomotive force (mmf) of the coil is $0.4\pi N_c I_c\sqrt{2}$ and this is consumed entirely inside the coil in the distance l_c. The field intensity is therefore

$$H_o = \frac{0.4\pi N_c I_c \sqrt{2}}{l_c} \qquad (7)$$

This equation can be transposed to give

$$I_c = \frac{H_o l_c}{0.4\pi N_c \sqrt{2}} \qquad (8)$$

The total flux linking the coil turns is obtained by adding equations 1, 2, and 5.

$$\phi_0 = H_o\left\{\left(A_g + \mu A_w P + \frac{k_r p_c \delta_c}{2}\right) - j\left(\mu A_w Q + \frac{k_r p_c \delta_c}{2}\right)\right\} \qquad (9)$$

The coil voltage which must be applied to establish the field intensity H_o in the air gap is

$$\dot{E}_c = j\sqrt{2}\pi f N_c \phi_0 10^{-8} \qquad (10)$$

or

$$\dot{E}_c = \sqrt{2}\pi f N_c H_o 10^{-8}\left\{\left(\mu A_w Q + \frac{k_r p_c \delta_c}{2}\right) + j\left(A_g + \mu A_w P + \frac{k_r p_c \delta_c}{2}\right)\right\} \qquad (11)$$

The impedance of the long coil is found by dividing E_c, equation 11, by I_c, equation 8.

$$\dot{Z}_c = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}\left\{\left(\mu A_w Q + \frac{k_r p_c \delta_c}{2}\right) + j\left(A_g + \mu A_w P + \frac{k_r p_c \delta_c}{2}\right)\right\} \qquad (12)$$

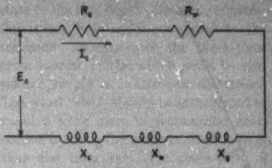


Fig. 2. Equivalent circuit, long coil

Recognizing the individual terms one can write,

$$\dot{Z}_c = (R_w + R_c) + j(X_g + X_w + X_c) \qquad (13)$$

where R_w is reflected work resistance, R_c is the resistance of the coil, X_g is reactance due to air-gap flux, X_w is reactance due to work flux, and X_c is reactance due to copper flux. Explicit expressions for these quantities are given in the following. The corresponding equivalent circuit for the long coil is shown in Fig. 2.

$$R_w = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}(\mu A_w Q) \qquad (14)$$

$$R_c = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}\left(\frac{k_r p_c \delta_c}{2}\right) \qquad (15)$$

$$X_g = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}(A_g) \qquad (16)$$

$$X_w = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}(\mu A_w P) \qquad (17)$$

$$X_c = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c}\left(\frac{k_r p_c \delta_c}{2}\right) \qquad (18)$$

The reactance X_c is often small in comparison to the total reactance in high-frequency jobs and may be neglected. It can, however, be quite important in close-coupled 60-cycle jobs.

By way of illustration, the expression for coil resistance will be derived by a separate line of reasoning. Consider an ideal single-turn coil consisting of a cylinder having internal perimeter p_c and length, l_c, and assume that current flows uniformly to depth, δ_c on the inside surface. The resistance to the flow of current is

$$R_1 = \frac{\rho_c p_c}{l_c \delta_c} \qquad (19)$$

According to the concept of effective depth of current penetration, δ_c, in equation 19, can be calculated by equation 6 if the radius of the cylinder is large compared to the calculated value of δ_c. Equation 19 therefore gives approximately the resistance of an ideal single-turn coil. Imagine that a narrow spiral cut is made to convert the cylinder into a coil of N turns. Its resistance will be increased by the factor

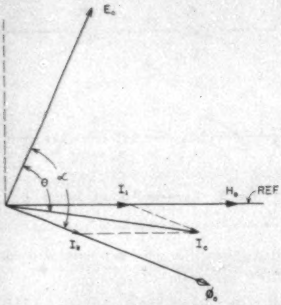


Fig. 3. Phasor diagram, general case (short coil)

N^2. But practical coils have appreciable spacing between turns and other imperfections so one would expect a practical coil to have a resistance

$$R_c = k_r \frac{\rho_c p_c N^2}{l_c \delta_c} \tag{20}$$

where k_r takes into account the spacing between turns and other imperfections of a practical coil. By using the relationship of equation 6 it can be shown that equation 20 is identical with equation 15.

The correction constant k_r is usually found to be slightly greater than unity as one would expect. For single-layer coils with a radial copper depth greater than δ_c, a good average value is $k_r = 1.15$.

The complex volt-amperes of the coil are calculated by multiplying E_c, equation 11, by I_c, equation 8.

$$\dot{E}_c I_c = 2.5 f H_o^2 l_c 10^{-9} \left\{ \left(\mu A_w Q + \frac{k_r p_c \delta_c}{2} \right) + j\left(A_g + \mu A_w P + \frac{k_r p_c \delta_c}{2} \right) \right\} \tag{21}$$

The real term in this expression is the sum of power input to the work and the power loss in the coil. The j term is the sum of the reactive volt-amperes due to air-gap flux, work flux, and copper flux. It is desirable to rewrite for future use the term corresponding to power input to the work.

$$P_w = 2.5 f H_o^2 l_c 10^{-9} (\mu A_w Q) \tag{22}$$

Since equation 22 gives the power input to work in terms of the magnetizing force H_o at the surface of the work and is not concerned with how much coil current is required to produce this H_o, it applies to short coils as well as to long coils. It is an important equation because it permits one to determine H_o and μ in terms of power density for substitution in equations 14 and 17. If the load is nonmagnetic, $\mu = 1$, and R_w and X_w may be calculated without regard to power density.

SHORT COILS

In short coils part of the total mmf is used up inside the coil and part outside the coil in the return flux path. The drop inside the coil is $H_o l_c$ and outside the coil is $R_e \phi_o$, where R_e is the reluctance of the external flux path. Thus

$$0.4\pi N_c \dot{I}_c \sqrt{2} = H_o l_c + R_e \phi_o \tag{23}$$

and

$$\dot{I}_c = \frac{1}{0.4\pi N_c \sqrt{2}} (H_o l_c + R_e \phi_o) \tag{24}$$

or

$$\dot{I}_c = I_1 + \dot{I}_2$$

The current I_1 flowing in the coil turns overcomes the internal drop and is in phase with H_o. This is the current which would flow if the reluctance of the external flux path were zero. The additional current I_2 supplies the external drop and is in phase with ϕ_o. Skin-effect considerations lead one to the conclusion that I_2 will not generally flow in the same path with I_1 but will tend to flow more on the back of the coil turns. It has been decided however to make the pessimistic and simplifying assumption in this paper that I_1 and I_2 flow through a common resistance and that this resistance is expressed by equation 15.

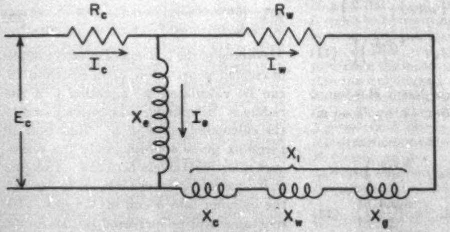


Fig. 4. Equivalent circuit, short coil

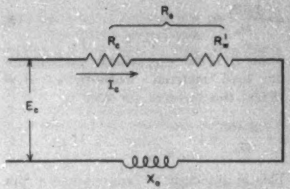


Fig. 5. Reduced equivalent circuit, short coil

The phasor diagram for the short coil is shown in Fig. 3. The extra current I_2 is shown in phase with the total flux ϕ_o, and the coil voltage is drawn approximately 90 degrees ahead of ϕ_o. Due to the fact that a small current-resistance drop results from the added current I_2 the coil voltage E_c actually leads ϕ_o by an angle α slightly less than 90 degrees.

The equivalent circuit diagram for the short coil is obtained by adding a shunt reactance X_e to the diagram for the long coil as shown in Fig. 4. The value of this reactance is obtained by dividing the coil voltage E_c, equation 10, by I_2 from equation 24. Thus

$$X_e = \frac{8\pi^2 f N_c 10^{-9}}{R_e} \tag{25}$$

To make this equation conform with equations 14–18 used to calculate the other circuit constants, this may be rewritten

$$X_o = \frac{8\pi^2 f N_c^2 10^{-9}}{l_c} \left(\frac{l_c}{R_e} \right) \tag{26}$$

The circuit diagram of Fig. 4 can be further reduced by solving for the equivalent resistance and reactance of the two parallel branches. This reduces the diagram to that shown in Fig. 5. Making the substitution

$$X_1 = (X_g + X_w + X_c) \tag{27}$$

one obtains the following expressions for R_w' and X_o.

$$R_w' = \frac{R_w X_e^2}{R_w^2 + (X_1 + X_e)^2} \tag{28}$$

$$X_o = \frac{X_e (R_w^2 + X_1^2 + X_e X_1)}{R_w^2 + (X_1 + X_e)^2} \tag{29}$$

The total resistance in the circuit is R_o, the sum of the coil resistance R_c and the reflected work resistance R_w'. The total reactance is X_o. One can now write

$$Z_o = \sqrt{R_o^2 + X_o^2} \tag{30}$$

$$I_o = \frac{E_c}{Z_o} \tag{31}$$

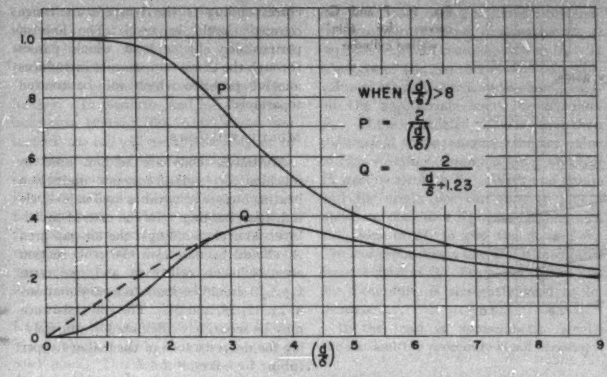


Fig. 6. P and Q curves for solid cylinder

$$\text{Coil power factor} = \frac{R_o}{Z_o} \qquad (32)$$

$$\text{Coil efficiency} = \frac{R_w'}{R_o} \qquad (33)$$

EXTERNAL RELUCTANCE R_e

It is necessary to find a way to calculate R_e. An accurate calculation would be very difficult and is not justified. On the basis of the change in inductance of a cylindrical solenoid with length, it is possible to show that its external reluctance is practically independent of length and equal approximately to $0.57/d$ where d is the coil diameter in centimeters. It is reasonable to assume that R_e for any coil shape is approximately inversely proportional to perimeter. Therefore it will be assumed that R_e may be calculated quite generally by

$$R_e = k_e \frac{1.80}{p_c} \qquad (34)$$

where k_e is an empirical factor to be evaluated. Because of lack of sufficient data to determine a better value of k_e, it will be assumed that $k_e = 1.0$.

SHORT COIL REPRESENTS THE GENERAL CASE

The short coil is obviously the general case because the equivalent circuit and all equations for the short coil degenerate into those for the long coil when the external reluctance R_e becomes negligible in comparison to the internal reluctance. The amount of end effect (short coil effect) can probably be best determined by a comparison of I_1 and I_2 in equation 24. If I_2 is negligibly small compared to I_1 the end effect is negligible.

P AND Q CURVES

The dimensionless flux factors P and Q are defined by equation 2. If one can calculate the total flux in a work piece of given cross section when a field intensity H_o exists at its surface, equation 2 can obviously be solved for P and Q. In this manner the P and Q curves for the solid cylinder and slab, Figs. 6 and 7, were determined. As would be expected, the values of P and Q depend not only on geometry but also on frequency, resistivity, and permeability. Since the depth of current penetration (equation 6) depends in the same way on these same quantities, it is convenient to plot the P and Q curves against the ratios d/δ and t/δ, respectively. This calculated δ is very close to the true depth of current penetration when it is small compared to the minimum dimension of the work-piece cross section. Otherwise it must be considered simply as a reference dimension. The word "slab" should be interpreted to mean a work piece of rectangular cross section in which the width is much greater than the thickness. The derivation of the P and Q curves for these two cases may be found in references 7 and 8.

The dotted portion of the Q curves of Figs. 6 and 7 represent empirical corrections to the theoretical curves which should be used when the load is magnetic. This correction is based on the findings of reference 1 (see appendix).

P and Q curves for a thin walled nonmagnetic hollow cylinder are given in Fig. 8. These were derived on the basis of transformer theory assuming uniform current distribution throughout the wall thickness and are consequently subject to certain limitations. These curves may be used with reasonable accuracy when

$$\frac{t}{d} \leq 0.10$$

and

$$\frac{t}{\delta} \leq 0.20$$

The symbol δ has even less physical significance in this case but it still is a good reference dimension.

Accurate P and Q curves for a nonmagnetic hollow cylinder are shown in Figs. 9 and 10. These were derived from the solution of the standard differential equation for the solid or hollow cylinder. The introduction of boundary conditions

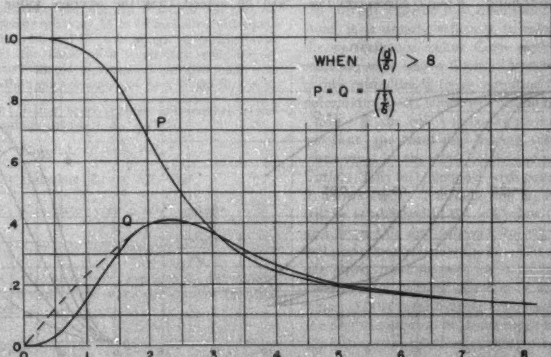


Fig. 7. P and Q curves for slab

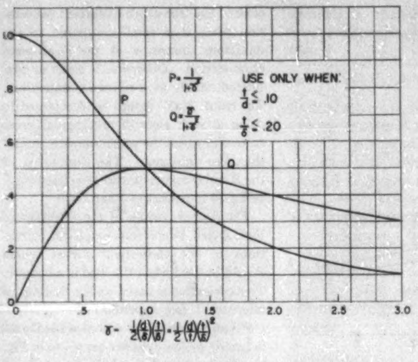


Fig. 8. P and Q curves for thin-walled cylinder

for the hollow cylinder considerably complicate the solution, and the calculation of values for plotting P and Q curves is extremely tedious. The solution can be used to calculate P and Q curves for other values of μ but this has not been done. Although the author has not been able to find a separate confirmation of this solution, it is apparently correct. The curves of Figs. 9 and 10 agree at one extreme ($t/d = 0.50$) with the solid cylinder curves of Fig. 6 and at the other extreme, where t/d and t/δ are both very small, with the thin-walled P and Q curves of Fig. 8.

Whereas the P and Q curves included here should suffice for most practical loads, there is nothing to prevent the addition of other P and Q curves as necessary or desirable. Curves may be determined either by calculation or by experiment.

Some ingenuity must obviously be used in choosing the proper P and Q curves for the job at hand. P and Q curves could be determined for pieces of square or rectangular cross section but this has not been done. It is suggested that a piece with a rectangular cross section which is square or nearly square be treated as a cylinder of equal perimeter. If the width of the rectangular cross section is equal to or greater than twice its height, the piece should be treated as a slab. A hollow cylinder may be treated as a solid cylinder if the wall thickness is greater than the depth of current penetration, equation 6. Regular shapes for which P and Q curves are given need no explanation. However, a thin strip or sheet is characteristic of a slab and a fine wire or rod is characteristic of a cylinder.

RESISTIVITY

Practically all induction-heating processes involve heating through a temperature range. Since it is impractical to take into account the varying resistivity, it will be agreed that the average value corresponding to the temperature range covered should be used. The loss of permeability by a load which passes through the Curie temperature introduces another problem which will be treated separately.

MULTILAYER COILS

Multilayer coils are seldom used in machine or radio-frequency induction heating but are sometimes used in 60-cycle induction heating. In the case of multilayer 60-cycle windings, the air-gap area A_g should be based on the cross-section area of mean coil turn and the term $k_r p_c \delta_c/2$ should be made zero in equations 9, 11, 12, 18, and 21. The coil resistance may be separately calculated by multiplying the d-c resistance of the hollow copper tubing by a factor of 1.5.

Application of Equations

Now that certain concepts and equations have been established, it will be shown how these may be used to design practical coils.

The problem usually is to design a coil to do a certain job and predict its performance. The equations previously developed allow one to follow either of two procedures in determining the desired results: the "classical procedure" which involves fundamental relationships of flux, voltage, and current; and the "equivalent circuit procedure" which employs the concept of an equivalent circuit which can be solved by conventional methods. For all practical purposes the two are equivalent and may be used interchangeably. They are presented in parallel in this paper. The classical procedure seems more convenient for the straightforward calculation of a magnetic load, but the

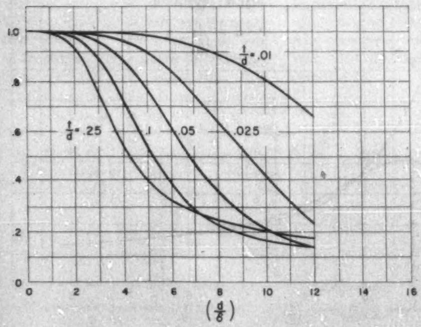


Fig. 9. Accurate P curves for hollow nonmagnetic cylinder, $\mu=1$

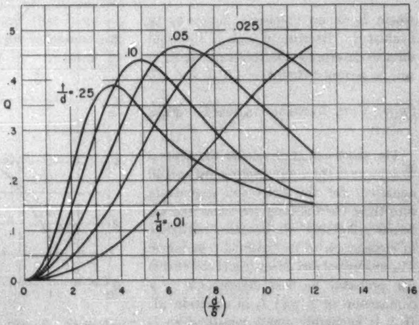


Fig. 10. Accurate Q curves for hollow nonmagnetic cylinder, $\mu=1$

equivalent circuit concept is often more useful in special situations. The choice of which to use is largely a matter of personal preference.

The two procedures may be subdivided further. In medium- and low-frequency induction heating (60–10,000 cycles per second) the coil will usually be designed to operate at a certain fixed voltage, whereas in radio-frequency heating (100–500 kc) it will usually be designed for a certain fixed current. This is characteristic of the generators used in the two frequency ranges. The two calculating procedures vary depending on whether the design is for fixed voltage or fixed current.

CLASSICAL PROCEDURE: VOLTAGE FIXED

1. Determine total power, P_w, required in work piece. This depends on the job to be done and must be separately determined before a coil calculation is started. It includes not only thermal power to raise the temperature of a certain mass of metal at the desired rate, but also conduction, convection, and radiation losses, heats of vaporization, heats of fusion, etc., when these are significant.

2. Determine H_o and μ from equation 22.

3. Calculate the absolute value of total flux from equation 9.

4. Choose coil voltage and substitute its value with the absolute value of ϕ_o in equation 10 to determine the number of coil turns. This is the most critical stage in fixing the coil design because at this stage one decides on the physical structure of the coil and whether or not a transformer will be required. The coil may be single turn or multiturn. It may consist of several multiturn sections in parallel or a multilayer coil of hollow copper tubing may be necessary to operate at standard line voltage as, on occasion, in 60-cycle induction heating. It is interesting to note that the coil design is fixed at this point without need to consider end effect. The coil will automatically draw the correct amount of current to induce the desired power P_w in the work piece although one does not know at this point the magnitude of this current or its phase relationship to the applied voltage. The calculation of this current and other performance data requires that end effect be taken into account. This is accomplished automatically when coil current is calculated in step 5.

5. Calculate absolute value of coil current from equation 24.

6. Calculate coil resistance from equation 15.

7. Calculate coil loss $P_c = R_c I_c^2$.

8. Calculate coil input $P_o = P_c + P_w$.

9. Calculate coil volt-amperes $(VA)_c = E_c I_c$.

10. Calculate coil power factor $= \dfrac{P_o}{(VA)_c}$.

11. Calculate coil efficiency $= \dfrac{P_w}{P_o}$.

It should be noted that in calculating a magnetic load, step 2 requires a cut-and-try procedure because, in equation 22, Q depends on μ and consequently on H_o. It is suggested that one assume a reasonable value of H_o, calculate μ and δ_w, and pick Q off the appropriate curve to calculate P_w. If the calculated value is lower than the desired value, assume a higher value of H_o; if higher, assume a lower value of H_o and recalculate P_w. With two points and the origin one can roughly plot a section of the curve P_w versus H_o and pick the value of H_o to give the desired P_w. The corresponding value of μ is calculated from equation 3. From this point on, the procedure is straightforward as indicated.

If the load is nonmagnetic ($\mu = 1$), the cut-and-try procedure is not necessary.

CLASSICAL PROCEDURE: CURRENT FIXED

1. Repeat steps 1 and 2 to determine H_o and μ.

2. Use the available current from the generator and equation 24 to determine the number of coil turns required. This requires a separate calculation of R_s and ϕ_o for substititon in equation 24.

3. Calculate the absolute value of coil voltage from equation 10 using N_c and ϕ_o as determined in step 2.

4. Proceed with the calculation of other data as in steps 6 through 11 in the previous section. Power-factor-correcting capacitors will not be used because power-factor correction is inherent in a radio-frequency generator.

EQUIVALENT CIRCUIT PROCEDURE: VOLTAGE FIXED

1. Duplicate steps 1 and 2 of classical method to determine H_o and μ. Note that this is done in every case.

2. On the basis of a single-turn coil ($N_c = 1$), calculate R_w, R_c, X_g, X_w, X_c, X_s, and X_1 using equations 14 to 18 inclusive and equations 26 and 27. Although R_c is calculated on the basis of a single-turn coil, the k_r used should correspond to the final coil which may or may not be a single-turn coil.

3. Using these values and equations 28 and 29, calculate R_w' and X_o.

4. Calculate $R_o = R_c + R_w'$.

5. Calculate $Z_o = \sqrt{R_o^2 + X_o^2}$.

6. Calculate coil power factor $= \dfrac{R_o}{Z_o}$.

7. Calculate coil efficiency $= \dfrac{R_w'}{R_o}$.

8. Calculate volts required on this single-turn coil or the volts per turn of the final coil.

$$E_{c1} = Z_o \sqrt{\frac{P_w}{R_w'}}$$

9. Choose coil voltage and determine coil turns.

$$N_c = \frac{E_c}{E_{c1}}$$

10. Multiply Z_o for the single-turn coil by N_c^2 to obtain Z_o for the final coil and calculate coil current.

$$I_c = \frac{E_c}{Z_o}$$

11. Calculate coil volt-amperes from product $E_c I_c$ or from P_w and power factor.

EQUIVALENT CIRCUIT PROCEDURE: CURRENT FIXED

1. Duplicate steps 1 through 7 in preceding section. Note that magnetizing force H_o, coil power factor, and coil efficiency are the same regardless of whether the coil is designed to operate at fixed voltage or fixed current.

2. Using available current I_c, total power to work P_w, and reflected work resistance in the single-turn coil R_w', calculate coil turns.

$$N_c = \frac{1}{I_c} \sqrt{\frac{P_w}{R_w'}}$$

3. Multiply (Z_o) of the single-turn coil by N_c^2 to obtain Z_o for the final coil and calculate coil voltage.

$$E_c = I_c Z_o$$

4. Calculate coil volt-amperes from product $E_c I_c$.

COIL TO HEAT THROUGH THE CURIE TEMPERATURE

If a magnetic work piece is to be heated from some temperature below Curie to some temperature above Curie, the calculation must be made in two steps. It is advisable first to design the coil to put a certain power into the work above Curie and determine its power-factor, efficiency, etc. It is then necessary to determine what this coil will do below Curie. If the current is to be held essentially constant, it is usually sufficient to calculate the performance below Curie with the same H_o used above Curie. A more accurate procedure is to use equation 24 to determine the H_o which will obtain below Curie. If the coil voltage is to remain constant, the total flux linking the coil turns must be the same below and above Curie. This fact, coupled with equation 9, allows one to determine the H_o which will be established in the coil below Curie and from this one proceeds to calculate coil performance below Curie. The average power must be such as to produce a certain total heating effect in a certain time. If the total heating time based on the first assumption of power input above Curie is not that desired, another assumption of power input above Curie must be made and the procedure repeated until the desired heating time is obtained.

WORK PIECE SHORTER THAN COIL

For clarity of exposition, it has been tacitly assumed that the work piece is at least as long as the coil. In practical loads, and especially in stationary heating, the work piece will usually be somewhat shorter than the coil. When this condition exists one should replace l_c in equation 22 by l_w, the length of the work piece. The general rule is to use either l_c or l_w, whichever is smaller. If one uses the equivalent circuit method and the work piece is shorter than the coil, the reflected work resistance R_w calculated from equation 14 should be reduced in the ratio l_w/l_c. Accuracy will be sacrificed in any case if the work piece is less than 80 per cent as long as the coil.

Choice of Frequency

Although it is really outside the scope of this paper, a word should be said about the choice of frequency for induction heating.

It is generally advisable to use the lowest frequency which will do the job with reasonable efficiency. This is where experience both in calculating and practical application is most important because many factors must be integrated in the decision. A good rule in the case of through heating is to choose a frequency which makes the depth of current penetration calculated from equation 6 one third to one fifth the minimum cross-section dimension of the work piece. This results in optimum power factor and efficiency. Note that the depth of current penetration in magnetic materials depends on μ which, in turn, depends on power density. Thus, a thin magnetic strip might be heated quite efficiently at 10,000 cycles per second at some low power density whereas the efficiency would become prohibitively low if one tried to obtain a much higher power density. In the surface hardening of steel, frequency should be chosen high enough to make the depth of current penetration above Curie less than about three times the maximum depth of hardening desired. A sufficient power density must also be provided at this frequency (short heating time) to prevent undue heat flow into the piece by conduction.

The standard frequencies available for induction heating are 60 cycles (commercial power lines), 960 cycles, 3,000 cycles, 9,600 cycles, and 450 kc. After deciding on the theoretically best frequency for a job, one usually chooses the next higher standard frequency unless economies dictate the choice of the next lower standard frequency with some sacrifice in efficiency. Because of the relatively low first cost of 60-cycle induction-heating equipment, one is often justified in sacrificing some efficiency to use this frequency. When heating through Curie as in heating billets for forging, 60-cycle induction may be used below Curie with some higher frequency used to finish the heating above Curie. This is called dual-frequency heating. The first cost of equipment rises steadily as frequency increases, but the ability to do a job and operating cost must be carefully considered in deciding on the frequency to use.

Conclusions

This method, in essentially its present form, has been spot-checked at all frequencies normally used for induction heating and its accuracy appears to be sufficiently good for general application and negotiation work. This does not mean that one can design a coil, build it, and be sure that it will be exactly right. It means that the calculated coil will be approximately right for the job and only a little tailoring will be necessary in the laboratory or on the job.

Those negotiating and doing application work will appreciate that this kind of accuracy is usually sufficient. Better accuracy may later be achieved by finding better values for k_r and k_e or possibly by introducing one or two additional factors which can be adjusted in accordance with experience. It should be pointed out that the equations take into account at least approximately all variable factors such as frequency, load shape, and material, and that the empirical factors k_r and k_e and any others, such as a reactance factor k_x which might be added, are really second-order corrections. The value of any of these factors should not differ greatly from unity. This is entirely different from the situation which maintains when one uses a more empirical method and alters the empirical factors to account for changes in load conditions as well as for the much smaller errors resulting from simplifying assumptions.

As mentioned previously, other P and Q curves may be determined for special shapes if this proves desirable. These may be determined either mathematically or by experiment.

Centimeter units were used here to retain the familiar form of equations. For everyday use, one may wish to rearrange the equations and introduce conversion factors so that dimension may be substituted in inch units.

Appendix. Errata in Reference 1

Certain errors unfortunately crept into reference 1 as it appears in the AIEE *Transactions* for 1945.

Equation 2 should read:

$$W=\frac{16.6NC\theta}{l_c} \text{ watts per cubic centimeter}$$

Equation 3 should read:

$$W=1/2\mu H^2 fG(kt)10^{-7} \text{ watts per cubic centimeter}$$

The equation appearing just below Fig. 4 should read:

$$k=\sqrt{\frac{4\pi^2\mu f10^{-9}}{\rho}}$$

Equation 6 should read:

$$G(kt)=\frac{2W10^7}{\mu H^2 f}$$

Equation 13 should read:

$$(VI)_a=2.5K_L fH^2 h_c l_c b_c 10^{-8} \text{ volt-amperes}$$

There is also some confusion introduced by the use of upper-case K interchangeably with lower-case k. Both have the same meaning, and were intended to be lower case.

References

1. INDUCTION HEATING OF MOVING MAGNETIC STRIP, R. M. Baker. *AIEE Transactions*, vol. 64, 1945, pp. 184–89.

2. HEATING OF NONMAGNETIC ELECTRIC CONDUCTORS BY MAGNETIC INDUCTION–LONGITUDINAL FLUX, R. M. Baker. *Ibid.*, vol. 63, June 1944, pp. 273–78.

3. DESIGN OF INDUCTION-HEATING COILS FOR CYLINDRICAL NONMAGNETIC LOADS, J. T. Vaughan, J. W. Williamson. *Ibid.*, vol. 64, Aug. 1945, pp. 587–92.

4. DESIGN OF INDUCTION HEATING COILS FOR CYLINDRICAL MAGNETIC LOADS, J. T. Vaughan, J. W. Williamson. *Ibid.*, vol. 65, 1946, pp. 887–92.

5. INDUSTRIAL ELECTRONICS REFERENCE BOOK (book), D. Venable, T. P. Kinn. "Radio Frequency Heating," chap. 24. John Wiley & Sons, Inc., New York, N. Y., 1948.

6. RADIO FREQUENCY HEATING (book), George H. Brown, Cyril N. Hoyler, Randolph A. Bierworth. D. Van Nostrand Company Inc., Princeton, N. J., 1947.

7. TRANSIENT ELECTRICAL PHENOMENA AND OSCILLATIONS (book), C. P. Steinmetz. McGraw-Hill Book Company, Inc., New York, N. Y., 1920.

8. MATHEMATICS APPLIED TO ELECTRICAL ENGINEERING (book), A. G. Warren. Chapman and Hall Ltd., London, England, chap. XVIII, 1941.

9. CALCULATIONS FOR CORELESS INDUCTION FURNACES, H. B. Dwight, M. M. Bogal. *AIEE Transactions (Electrical Engineering)*, vol. 54, 1935, pp. 312–16.

(Discussion follows)

Discussion

C. A. Tudbury (TOCCO Division, Ohio Crankshaft Company, Cleveland, Ohio): Mr. Baker's opening paragraphs suggest that he has observed reticence on the part of engineers to use a quantitative approach to the problem of induction-heating-coil calculation and design. We have made the same observation. The simplicity of Mr. Baker's method of calculation should prove helpful in overcoming this.

In discussing his P and Q curves for nonmagnetic hollow cylinders, he mentions that he has not been able to find separate confirmation of his solution. We have also calculated and plotted numerical results from the solution of the differential equation for round hollow loads. Our equation is in a slightly different form. Substituting

$$A_w=\frac{\pi d^2}{4}$$

and

$$\delta=\sqrt{\frac{\rho}{4\pi^2\mu f 10^{-9}}}$$

into equations 14 and 17 yields

$$Z_w=\frac{\rho\pi d}{\delta l_c}\left[\frac{d}{2\delta}\,Q+j\,\frac{d}{2\delta}P\right]N_c^2=\frac{\rho\pi d}{\delta l_c}[K_{R3}+jK_{X3}]\,N_c^2$$

The first part of this expression, outside the brackets, is in the form of the familiar relationship for the resistance of a conductor-resistivity times length divided by area. Length is the outside periphery of the work in the direction of induced current flow. Area is the depth of current penetration (equation 6) multiplied by coil length. The real and the imaginary factors, K_{R3} and K_{X3}, within the brackets, correspond to the author's Q and P, respectively, except that they contain a factor equal to one half the ratio of work-piece diameter to current penetration.

K_{R3} and K_{X3} are the real and imaginary components of the expression in brackets in the following:

$$Z_w=\frac{\rho\pi d}{\delta l_c}N_c^2\left[\frac{\sqrt{2}(ber'\,z_0+j\,bei'\,z_0)}{(1-g)(ber\,z_0+j\,bei\,z_0)}-\frac{g\sqrt{2}(ker'\,z_0+j\,kei'\,z_0)}{(1-g)(ker\,z_0+j\,kei\,z_0)}\right]$$

where

$$g=\left[\frac{ker\,z_0+j\,kei\,z_0}{ber\,z_0+j\,bei\,z_0}\right]\times\left[\frac{\left(ber\,z_1-\frac{2u}{z_1}bei'\,z_1\right)+j\left(bei\,z_1+\frac{2u}{z_1}ber'\,z_1\right)}{\left(ker\,z_1-\frac{2u}{z_1}kei'\,z_1\right)+j\left(kei\,z_1+\frac{2u}{z_1}ker'\,z_1\right)}\right]$$

and

$$z_0=\frac{d}{\sqrt{2}\,\delta};\quad z_1=\frac{d_1}{\sqrt{2}\,\delta};\quad d_1=d-2t$$

This agrees with equation 49 from reference 3 of the paper. The numerical values given in this equation in (K_3 and K_4) presuppose a large ratio of work-piece outside diameter to current penetration. Our calculations of K_{R3} and K_{X3} do not depend upon such an assumption. They apply for the full range of work-piece wall thickness, from zero wall to a solid bar. When the factor $d/2\delta$ in equation 3 is taken into account, we find excellent numerical agreement between our curves and Mr. Baker's. This is particularly gratifying because, as pointed out in the paper, the calculations leading to these results are tedious.

Paul P. Biringer (University of Toronto, Toronto, Ont., Canada): Mr. Baker has done an excellent piece of work in presenting a simplified theory for designing induction-heating coils. It is well known that great difficulties are encountered in obtaining accurate methods of design, mainly because of the complexity involved in solving the differential equations of the heat flow and magnetic field simultaneously.

Mr. Baker's method is, in general, very useful; however some caution must be taken in using equations 3 and 4 for magnetic loads. In the case of surface hardening these equations and the method described do not apply. From the theoretical point of view, both equations 1 and 6 are valid only for homogeneous materials. In the case of surface hardening and during the heating period for magnetic materials, the charge is far from homogeneous. In the latter case one has to deal with a red-hot nonmagnetic surface layer and a relatively cold magnetic core and they are both conducting eddy currents. Experience has shown that under such circumstances the effect of eddy currents is not covered by the theory which is generally used. Equation 2 assumes that the entire eddy-current sheet is accommodated in the nonmagnetic layer. This is true only if the core can be assumed devoid of flux and current. Dr. Dreyfus[1] found, however, that often a considerable portion of the eddy-current sheet will fall within the magnetic zone of the charge. In the case mentioned here, the core will carry a flux so great that the eddy-current density will diminish at a much lower rate than indicated by the exponential relationship ($e^{-(1+j)x/\delta}$) on which equation 2 was based. The surface power density will be much higher than would be expected. A simple example can be given to illustrate this. Assume a thin nonmagnetic tube having a laminated magnetic core heated by induction. For the same current through the heating coil the surface power rises considerably. Generally speaking, when energy is dissipated in a magnetically zoned material the eddy-current density in the outer nonmagnetic zone diminishes at a lower rate than indicated by the exponential function, and even during the final heating period the surface power is considerably greater than might be expected from the equations which are based on homogeneous materials.

It can be concluded that both equation 6 and the design procedure will give correct results if the material is homogeneous. There will be a negligible error if the magnetic material is below its Curie point and if varying permeability is assumed. Large errors can be expected in some cases however, if a nonmagnetic surface layer envelops a magnetic core. In the latter case one must be careful in defining the effective depth of current penetration. It can be stated that if the depth of the nonmagnetic layer is less than the depth of current penetration, the error will be considerable when using the conventional theory. If the depth of the nonmagnetic surface layer is less than 75 per cent of the depth of current penetration, the conventional theory cannot be used. The error would be so large in this latter case that one should work with a modified theory which takes into account the magnetic layer inside the nonmagnetic zone. The case hardening of crank shafts will often fall within this category.

There are some minor points where it is difficult to agree with the author. For example, it is not clear how skin effect can cause the concentration of current on the inside of the coil.

REFERENCE

1. HIGH FREQUENCY HEATING AND TEMPERATURE DISTRIBUTION IN SURFACE HARDENING OF STEEL, L. A. **Dreyfus**. *Acta Polytechnica*, Stockholm, Sweden, Electrical Engineering Series, vol. 4, no. 5.

H. R. Weed (The Ohio State University, Columbus, Ohio): The author is to be complimented on his work to simplify induction-heating-coil design. It is of interest to point out that the techniques outlined by Mr. Baker for calculation of induction-heating coils appear to be very similar to the methods discussed in at least two textbooks.[1,2]

The simplifications of constant permeability and small skin depth represent one special case of equation 10.63 of reference 1 and is expressed specifically in equation 10.65. The calculation of the real and reactive components of impedance are carried out with equal simplicity in equations 10.53 and 10.54 or equation 10.96 of reference 1, depending on the coil spacing. It is also felt that the plots of Figs. 10-12, 10-14, and 10-15 of reference 1 are more useful as an indication of the proper direction to take in design than the P and Q curves of this paper.

My experience indicates that the greatest single error in this approach is the assumption of constant permeability, particularly when working with ferrous materials.

REFERENCES

1. INDUSTRIAL ELECTRONIC ENGINEERING (book), W. L. **Davis**, H. R. **Weed**. Prentice-Hall, Inc. New York, N. Y., 1953.

2. Reference 6 of the paper.

R. M. Baker: The supplementary material supplied by Mr. Tudbury substantiates my statement that the calculation of P and Q curves for the hollow cylinder is tedious. It is gratifying that his results agree with those in the paper.

Mr. Biringer has chosen to discuss principally coil design for surface hardening. This is a special application of high-density heating where the surface layer is heated to hardening temperature with the remainder of the piece remaining relatively

Table I

	Test	Calculation
Coil impedance (Z_o)	1.0×10^{-3} ohms	0.82×10^{-3}
Coil resistance (R_o)	0.547×10^{-3} ohms	0.313×10^{-3}
Coil reactance (X_o)	0.830×10^{-3} ohms	0.75×10^{-3}
Coil power factor	0.547	0.372
Coil efficiency	—	72 per cent

cool. He is correct in stating that the method of this paper is not applicable in the special situation where the surface layer has gone above Curie and where the thickness of this layer is still considerably less than the effective depth of current penetration calculated for above Curie conditions. I did not give much thought to hardening coils in setting up the calculating procedures of this paper, and in fact most hardening coils would be ruled out by the more or less arbitrary limitation placed on the ratio of length to inside diameter of coil. More empirical procedures are usually employed in designing and applying hardening coils, although one lacking this background experience and information would receive considerable help and guidance through applying the methods of this paper. Because magnetizing forces used for induction hardening are high, the effective magnetic permeability is low. This, coupled with the fact that a surface layer of considerable depth is above Curie during most of the heating cycle, leads one to expect that the coil constants will be determined to a considerable extent by the above Curie properties of the load. That this is indeed true is illustrated in Table I where the measured constants of a 3,000-cycle induction-hardening coil are compared with calculated constants based on above Curie conditions ($\rho=110\times10^{-6}$ ohm-centimeters and $\mu=1$). The coil has an inside diameter of $2^{11}/_{16}$ inches and a length of $1^{3}/_{4}$ inches. It was used to harden progressively a $2^{3}/_{8}$-inch-outside-diameter shaft at a rate of 0.30 inch per second to a depth of 0.20 inch.

Whereas this would certainly not be considered a satisfactory check between test and calculated results, it does serve to emphasize the fact that the magnetic effect is not large in progressive induction hardening. In the present case at least, which is fairly characteristic of deep case hardening, the magnetic contribution increases effective coil resistance by a factor less than two and increases coil reactance and impedance by only 10–20 per cent.

Mr. Weed points out a similarity between certain equations in the subject paper and those in his reference 1. This is understandable because both are based on standard skin-effect equations. I cannot agree if Mr. Weed means to indicate that the two treatments will produce comparable results in the hands of the average user. In his reference the treatment is essentially limited to the solid cylinder without specific instructions on how the user is to take into account air gap or the end effect of short coils. Also, no specific instruction is given on how one is to treat magnetic loads. In the present paper, I have tried to provide the user with a complete general calculating procedure applicable to loads of various shapes, both magnetic and nonmagnetic. For the person constantly involved in application work, this develops familiarity and facility. For the uninitiated or casual user it provides a method usually general enough to solve the problem at hand. Extreme accuracy is not usually required. The generality of application is obtained by assuming a general expression involving P and Q factors for the total work flux and by presenting curves of P and Q for various load shapes. The method is flexible because functions of P and Q for new configurations can be added from time to time as they are derived or determined experimentally. In using this method of calculation the user will soon discover that certain simplifications can sometimes be made in the equations without sacrificing accuracy. The equations can be very greatly simplified in all cases if one makes the assumption that the depth of current penetration is small compared to the minimum dimension of the work piece. I have carefully avoided indicating this simplification, however, because I have found people very prone to apply the simplified equations without regard for the limitation. This is in line with the general philosophy that a calculating method should automatically take into account at least approximately all factors which can appreciably affect results.

A New Stability Criterion of Linear Servomechanisms by a Graphical Method

TOSHIO NUMAKURA
NONMEMBER AIEE

T. MIURA
NONMEMBER AIEE

Synopsis: When the optimum condition of a control system is sought with the help of analog computers, the usual procedure is to vary parameters such as time constants and gains of control elements. However, with the analytical method, it is difficult to express the character of the system explicitly in terms of these parameters because some of the coefficients of the characteristic equation are functions of the parameters.

In this paper, a new graphical method, in which the stable region of the system is drawn on a plane with two of the parameters as co-ordinates, is proposed. This graphical method will be very useful in the analysis of systems with the help of analog computers. As the stability criterion given by this method using Hermite's theorem is absolutely equivalent to that by Hurwitz's theorem, this graphical method may be considered as a graphical representation of Hurwitz's theorem.

MANY METHODS for linear servomechanism analysis have been developed. However, with ordinary methods, when the characteristic equation of a system has several coefficients which are the functions of certain parameters, it is difficult to express its character explicitly in terms of these parameters. For example, with a system where the feedback transfer function of its minor loop contains several parameters, many coefficients of the characteristic equation also contain these parameters.

Since the end of World War II, Professor K. Tashiro has been studying methods of analyzing high-order systems, and his work provided the impetus for this paper. Another paper, published in the USSR, deals with the same problems but employs a different procedure. An abstract of this work has been noted.[1]

The following section discusses the system whose characteristic equation contains a single parameter in both its real and imaginary parts; the latter single parameter may be either the same as or different from the former. In this

---

Paper 57-185, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 9, 1956; made available for printing November 28, 1956.

Toshio Numakura and T. Miura are with Hitachi, Central Research Laboratory, Tokyo, Japan.

The authors wish to express their thanks to Professor K. Tashiro, Kyoto University, Doctor B. Tadano, and A. Kamoi for their advice and encouragement.

case, with these two parameters as co-ordinates, boundary curves can be drawn which divide the plane into stable regions and unstable regions. The stable and unstable regions are defined very easily by Hermite's theorem. Hereafter, this plane will be designated the *U-V* plane. As the stability criterion is absolutely equivalent to that of Hurwitz's theorem, this graphical method can be considered a graphical representation of this theorem.

Also considered is the system in which each coefficient of the characteristic equation depends linearly on the two parameters, *X* and *Y*. With these parameters as co-ordinates, boundary curves can be drawn, each of which separates the stable region from the unstable region. These boundary curves on the *X-Y* plane are peculiar to the system and show clearly the necessary relations between *X* and *Y* for the system to be stable. Moreover, these curves show the change of response frequency as well as that of the damping character caused by the variation in the parameters. The method is very practical and has been applied to analyses and syntheses of many systems. This paper analyzes systems which contain two parameters, that is, the gain and the time constant of the damping transformer for stabilization in the feedback minor loop.

U-V Plane

Boundary Curves and Stability Criterion on the U-V Plane

With a linear system, the characteristic equation, that is, the denominator of the closed-loop transfer function, is written

$$f(p)=\xi_n+\xi_{n-1}p+\xi_{n-2}p^2+\ldots+\xi_1p^{n-1}+\xi_0p^n \qquad (1)$$

In general, when the gain or the time constant of one control element is varied, one or several coefficients among ξ_n, ξ_{n-1}, ... , ξ_0 change simultaneously. In order for the system to sustain an oscillation when subjected to a step-function input, it is necessary for $f(i\omega)=0$. Putting $\omega=\sqrt{\gamma}>0$, then

$$\begin{aligned} f(i\omega)&=\xi_n+i\sqrt{\gamma}\,\xi_{n-1}-\gamma\xi_{n-2}-i\sqrt{\gamma}\gamma\xi_{n-3}+\gamma^2\xi_{n-4}-\ldots\\ &=(\xi_n-\gamma\xi_{n-2}+\gamma^2\xi_{n-4}-\ldots)+i\sqrt{\gamma}(\xi_{n-1}-\gamma\xi_{n-3}+\gamma^2\xi_{n-5}-\ldots)\end{aligned}$$

To satisfy the foregoing equation, both the real and the imaginary parts of the equation must be zero.

$$\left.\begin{aligned} U(\gamma)&=\xi_n-\gamma\xi_{n-2}+\gamma^2\xi_{n-4}-\ldots &&=0\\ V(\gamma)&=\xi_{n-1}-\gamma\xi_{n-3}+\gamma^2\xi_{n-5}-\ldots &&=0\end{aligned}\right\} \qquad (2)$$

Now, assume

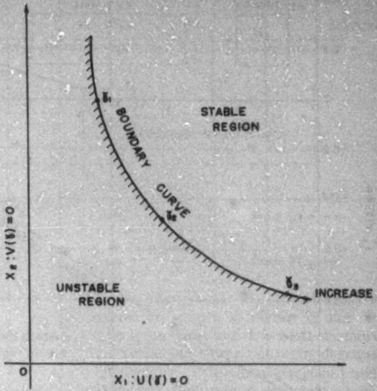


Fig. 1. Stability chart on the U-V plane

$$\xi_{n-2j}=a_{n-2j}X_1+b_{n-2j}\left(j=0,1,2,\ldots,\left[\frac{n}{2}\right]\right)$$

$$\xi_{n-(2j+1)}=a_{n-(2j+1)}X_2+b_{n-(2j+1)}\left(j=0,1,\ldots,\left[\frac{n-1}{2}\right]\right) \qquad (3)$$

From equations 2 and 3, the following is obtained

$$\left.\begin{aligned} X_1&=-\frac{\Sigma(-\gamma)^jb_{n-2j}}{\Sigma(-\gamma)^ja_{n-2j}}\\ X_2&=-\frac{\Sigma(-\gamma)^jb_{n-(2j+1)}}{\Sigma(-\gamma)^ja_{n-(2j+1)}}\end{aligned}\right\} \qquad (4)$$

corresponding to $U(\gamma)=0$ and $V(\gamma)=0$.

With parameters X_1 and X_2 as co-ordinates, curves can be plotted on the plane as shown in Fig. 1. This plane is the *U-V* plane, and the curves are the boundary curves. Furthermore, each value of $\sqrt{\gamma}/2\pi=f$ gives the frequency of continuous oscillation when the conditions are on the boundary curve. One side of each boundary curve is the stable region, and the other is the unstable. The stable and unstable regions are outlined by the following four conditions:

1. The number of roots of $U(\gamma)=0$ and $V(\gamma)=0$ are equal to the degree of these equations, respectively, and every root is positive.
2. The roots of $U(\gamma)=0$ and $V(\gamma)=0$ separate each other.
3. The smallest root belongs to $U(\gamma)=0$.
4. The coefficients of $U(\gamma)$ and $V(\gamma)$ have the same sign.

In the stable regions, all of these conditions are satisfied, whereas in the unstable regions one or more of them are not. It must be noted that, in many cases, neither side of a boundary curve satisfies these conditions. In that case, the system is unstable on both sides of and on the boundary curve; the boundary curve is not a boundary curve in the real sense of the word.

The foregoing conditions 1, 2, and 3 are

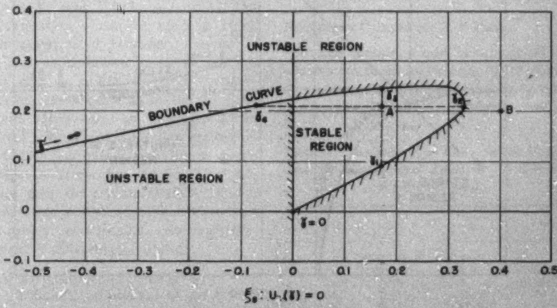


Fig. 2. Stability chart for equation 5, the case where a stable region exists

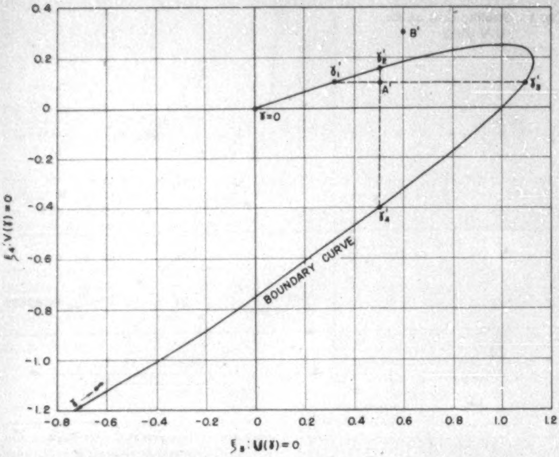


Fig. 3. Stability chart for equation 6, the case where no stable region exists

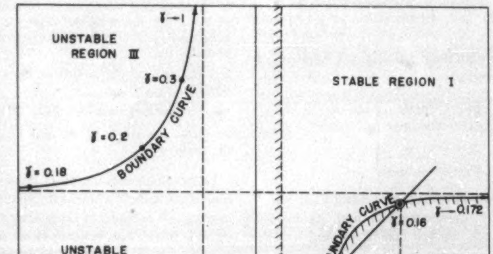


Fig. 4. Stability chart for equation 7

derived from Hermite's theorem. They indicate that all the roots of $f(p)=0$ exist on one side of the imaginary axis of the complex plane, but it is not known on which side these roots really exist. This means that all the coefficients of $U(\gamma)$ are of the same sign, and, similarly, that all the coefficients of $V(\gamma)$ are of the same sign, although it is not clear whether the two signs are same or not. Therefore, the fourth condition is necessary to show on which side of the complex plane these roots exist, or whether the sign of the coefficients of $U(\gamma)$ and $V(\gamma)$ are the same or not. In other words, if conditions 1, 2, and 3 are satisfied, and if any one coefficient of $U(\gamma)$ and any one coefficient of $V(\gamma)$ have the same sign, then the system is stable. Thus, with these four conditions, the stability criterion on the *U-V* plane becomes perfect.

In condition 2, the words "separate each other" mean that if the roots of $U(\gamma)=0$ and $V(\gamma)=0$ are arranged in order of magnitude, a root of *U*'s is formed between the adjoining two roots of *V*'s and conversely. Of course, it is also required that there are no multiple roots at all. The proof of these criteria are given in Appendix I; the method of appreciation is explained in detail in the following section, with simple examples.

STABILITY CHART OF (ξ_n, ξ_{n-1})

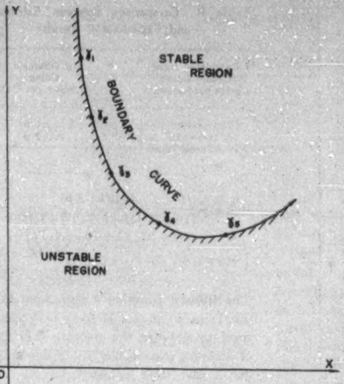
As the first example, consider a differential (or characteristic) equation

$$p^5+3p^4+p^3+2p^2+\xi_4 p+\xi_5=0 \qquad (5)$$

Putting $p^2=\gamma$, then

$$\xi_5=2\gamma-3\gamma^2 \text{ from } U(\gamma)=0$$
$$\xi_4=\gamma-\gamma^2 \text{ from } V(\gamma)=0$$

With the foregoing a boundary curve can be drawn as shown in Fig. 2. In this case, both $U(\gamma)$ and $V(\gamma)$ are of second degree, so each must have two positive roots according to condition 1. At point *A* in the first quadrant, the roots to satisfy $U(\gamma)=0$ are the cross points of the boundary curve and the straight line parallel to ξ_4 axis through *A*, that is, γ_1 and γ_3. The roots to satisfy $V(\gamma)=0$ are the cross points of the boundary curve and the straight line parallel to ξ_5 axis through *A*, that is, γ_2 and γ_4. Then, at point *A*, condition 1 is satisfied. As shown in Fig. 2, $\gamma_1<\gamma_2<\gamma_3<\gamma_4$, conditions 2 and 3 are satisfied; and both ξ_0 and ξ_1 are positive, so condition 4 is also satisfied. Point *A* is thus a stable point. On the other hand, at point *B*, there is no root to satisfy $U(\gamma)=0$; condition 1 is not satisfied, so it is an unstable point. In the second quadrant, there is only one root to satisfy $U(\gamma)=0$; in the third quadrant, one root



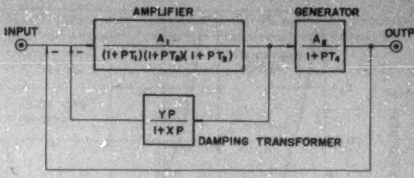


Fig. 5 (left). Stability chart on the X-Y plane

Fig. 6 (above). Block diagram of automatic voltage control system no. 1

for $U(\gamma)=0$ and $V(\gamma)=0$, respectively; and in the fourth quadrant, also one root for $V(\gamma)=0$. Thus the second, third, and fourth quadrants are all unstable regions. That is, the stable region exists only in the domain which is enclosed by the ξ_4 axis and the boundary curve in the first quadrant.

Next, consider

$$p^5+2p^4+p^3+3p^2+\xi_4 p+\xi_5=0 \qquad (6)$$

which obtains

$$\xi_5=3\gamma-2\gamma^2 \quad \text{from } U(\gamma)=0$$
$$\xi_4=\gamma-\gamma^2 \quad \text{from } V(\gamma)=0$$

With them the foregoing, a boundary curve can be drawn as shown in Fig. 3. In this case, both $U(\gamma)$ and $V(\gamma)$ are of second degree, so each must have two positive roots as in the first example. Point B' and the second, third, and fourth quadrants are all unstable regions for the same reason as in the first example. At point A' the roots of $U(\gamma)=0$ and $V(\gamma)=0$ are γ_2', γ_4', and γ_1', γ_3' respectively. In addition, as $\gamma_1'<\gamma_2'<\gamma_3'<\gamma_4'$ and every coefficient is positive, it satisfies conditions 1, 2, and 4. Therefore, it seems to be a stable point but the smallest root γ_1' belongs to $V(\gamma)=0$, so A' is also an unstable point. As a result, in the system of differential equation 6 there is no stable point for all the values of ξ_4 and ξ_5.

This fact is also proved easily by Hurwitz's theorem. If (0.5, 0.1) are chosen as the corresponding value to A', the following determinant is obtained:

$$\begin{vmatrix} 2 & 3 & \xi_5 & 0 \\ 1 & 1 & \xi_4 & 0 \\ 0 & 2 & 3 & \xi_5 \\ 0 & 1 & 1 & \xi_4 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 0.5 & 0 \\ 1 & 1 & 0.1 & 0 \\ 0 & 2 & 3 & 0.5 \\ 0 & 1 & 1 & 0.1 \end{vmatrix} = 0.11>0$$

But a principal minor

$$\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1<0$$

Thus, except for ξ_4 and ξ_5 even the coefficients do not satisfy Hurwitz's theorem, and so the system of differential equation 6 is unstable for any value of ξ_4 and ξ_5.

From the foregoing two examples, it follows that in the stability chart (ξ_n, ξ_{n-1}) there is a stable region when the vector from the origin to a point corresponding to a value of γ on the boundary curve rotates counterclockwise as γ increases, but there is no stable region when the vector rotates clockwise as γ increases. Any stability chart of (ξ_{n-2r}, $\xi_{n-(2r+1)}$) can be discussed similarly.

STABILITY CHART FOR A CASE WHERE MORE THAN THREE COEFFICENTS DEPEND LINEARLY ON A PARAMETER X

Consider a case where more than three coefficients depend linearly on a parameter X. For convenience, X is written in ξ_{n-2j} as X_1, and X in $\xi_{n-(2j+1)}$ as X_2, and thus equation 4 can be used as it is. With equation 4 boundary curves of (X_1, X_2) can be drawn for every value of γ as shown in Fig. 4. The stability criterion is the same as in the previous section. However, as $X_1=X_2=X$ in the real system, the values are limited on the 45-degree sloped straight line as shown in Fig. 4. If the line lies in the stable region then the system is stable, and if the line lies in the unstable region then the system is unstable. For example, consider a differential equation.

$$(p)=(X-0.0001)p^4+(1+4X)p^3+ \\ (0.003+6X)p^2+(0.06+4X(p+X \qquad (7)$$

From equations 4 and 7,

$$X_1=\frac{0.0001\gamma\,(30+\gamma)}{1-6\gamma+\gamma^2} \quad \text{from } U(\gamma)=0$$

$$X_2=\frac{\gamma-0.006}{4(1-\gamma)} \quad \text{from } V(\gamma)=0$$

From these equations the boundary curves are drawn as shown in Fig. 4. The upper-right and lower-left domains are stable and all the remaining domains are unstable. The unstable regions I and II do not satisfy at least one of the conditions 1, 2, and 3, and the unstable regions III and IV satisfy these three conditions but do not satisfy condition 4.

The straight line $X_1=X_2=X$ passes through the stable and unstable regions alternately. That is, the system is stable for $X<-0.3$, unstable for $-0.3<X<0.0001$, stable for $0.0001<X<0.00029$, unstable for $0.00029<X<0.033$, and stable for $0.033<X$.

On the other hand, the values calculated by Hurwitz's theorem are 0.0001, −0.3, 0.00029, and 0.033. The last three are calculated from $64X^3+17.0X^2-0.631X+0.00018=0$. These results completely coincide with those obtained on the stability chart. Therefore, according to the foregoing examples, the stability criterion on the U-V plane can be considered as a graphical representation of Hurwitz's theorem.

In Fig. 4, as the scale of co-ordinates is arbitrary, no scale designations are shown.

X-Y Plane

BOUNDARY CURVES ON X-Y PLANE

In this section, a case is considered in which some or all coefficients of the characteristic equation depend linearly on two parameters, X and Y. Assume

Table I. Coefficients of Characteristic Equation Calculated from Equation 15 (where $\xi_j=a_jX+b_jY+c_j$)

j	ξ_j	a	b	c
0	ξ_0	0.00342	0	0
1	ξ_1	0.0918	0	0.00342
2	ξ_2	0.75	0	0.0918
3	ξ_3	2.11	46.9	0.75
4	ξ_4	83.8	27.6	2.11
5	ξ_5	0	0	83.8

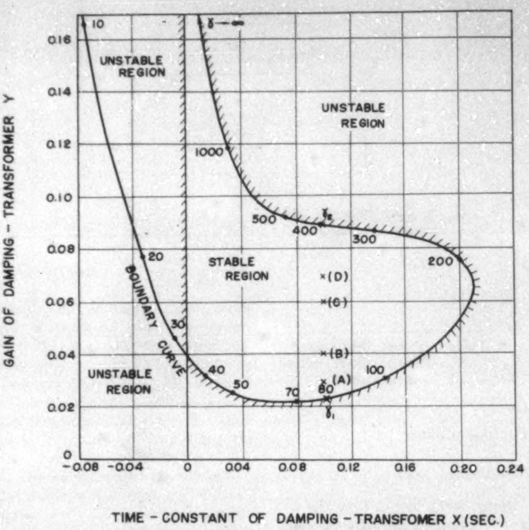


Fig. 7. Stability chart for automatic voltage control system no. 1

$$\xi_j = a_jX + b_jY + c_j \; (j=0, 1, \ldots, n) \tag{8}$$

where a_j, b_j, c_j are constants. From equation 2 is obtained

$$\left.\begin{aligned} U(\gamma) &= X\,\Sigma(-\gamma)^j a_{n-2j} + Y \\ &\quad \Sigma(-\gamma)^j b_{n-2j} + \Sigma(-\gamma)^j c_{n-2j} = 0 \\ V(\gamma) &= X\,\Sigma(-\gamma)^j a_{n-(2j+1)} + \\ &\quad Y\,\Sigma(-\gamma)^j b_{n-(2j+1)} + \\ &\qquad \Sigma(-\gamma)^j c_{n-(2j+1)} = 0 \end{aligned}\right\} \tag{9}$$

By solving these simultaneous equations,

$$\left.\begin{aligned} X &= \frac{\Sigma(-\gamma)^j b_{n-2j}\Sigma(-\gamma)^j c_{n-(2j+1)} - \Sigma(-\gamma)^j b_{n-(2j+1)}\Sigma(-\gamma)^j c_{n-2j}}{\Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)} - \Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j}} \\ Y &= \frac{\Sigma(-\gamma)^j c_{n-2j}\Sigma(-\gamma)^j a_{n-(2j+1)} - \Sigma(-\gamma)^j c_{n-(2j+1)}\Sigma(-\gamma)^j a_{n-2j}}{\Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)} - \Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j}} \end{aligned}\right\} \tag{10}$$

By use of equation 10, boundary curves are drawn on the X-Y plane as shown in Fig. 5. One side of each boundary curve is a stable region and the other side an unstable region as in the case of the U-V plane. These regions are discriminated by the following criteria (Appendix II).

If the denominator of X and Y

$$\Delta(\gamma) = \Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)} - \Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j} \tag{11}$$

is positive (negative) for a value of γ, then the left- (right-) hand side of the boundary curve on which the point corresponding to the γ lies is stable and the right- (left-) hand side is unstable, facing the direction in which γ increases.

The unstable regions differentiated by the foregoing criteria are always unstable. However, the stable regions differentiated by these criteria are not always stable. Therefore, Hermite's theorem or Hurwitz's theorem must be applied to these regions for exact discrimination; the foregoing criteria are not perfect, but are convenient for rough discrimination.

EXAMPLE: AUTOMATIC VOLTAGE CONTROL SYSTEM (No. 1)

The closed-loop transfer function of the system, Fig. 6, is given by

Table II. Comparison Between Calculated and Experimental Results

f Calculated from Boundary Curve, Cycles per Second	f Obtained by Computer, Cycles per Second
1.4(γ_1=80)	1.5
3.1(γ_2=380)	2.9

$$AG(p) = \frac{A_1A_2(1+Xp)}{(1+pT_1)\{(1+pT_1)(1+pT_2)\times(1+pT_3)(1+Xp)+A_1Yp\}+A_1A_2(1+Xp)} \tag{12}$$

The transfer function of the minor loop, $Yp/(1+Xp)$, is of a damping transformer used to stabilize the system. The gain of damping transformer will be designated as Y the time constant as X. The characteristic equation of the system is

$$\begin{aligned} &\Pi T_iXp^5 + (\Pi T_i + \Sigma T_iT_jT_k\cdot X)p^4 + \\ &\quad (\Sigma T_iT_jT_k + \Sigma T_iT_j\cdot X)p^3 + \\ &\quad (\Sigma T_iT_j + \Sigma T_i\cdot X + A_1T_4Y)p^2 + \\ &\quad \{\Sigma T_i + (1+A_1A_2)X + A_1Y\}p + \\ &\qquad (1+A_1A_2) = 0 \end{aligned} \tag{13}$$

where

$$\left.\begin{aligned} &\Pi T_i = T_1T_2T_3T_4 \\ &\Sigma T_iT_jT_k = T_1T_2T_3 + T_1T_2T_4 + \\ &\qquad T_1T_3T_4 + T_2T_3T_4 \\ &\Sigma T_iT_j = T_1T_2 + T_1T_3 + T_1T_4 + \\ &T_2T_3 + T_2T_4 + T_3T_4 \\ &\Sigma T_i = T_1 + T_2 + T_3 + T_4 \end{aligned}\right\} \tag{14}$$

Comparing of equation 13 with equation 1 obtains $n=5$ and

$$\left.\begin{aligned} \xi_0 &= \Pi T_i\cdot X \\ \xi_1 &= \Sigma T_iT_jT_k\cdot X + \Pi T_i \\ \xi_2 &= \Sigma T_iT_j\cdot X + \Sigma T_iT_jT_k \\ \xi_3 &= \Sigma T_i\cdot X + A_1T_4Y + \Sigma T_iT_j \\ \xi_4 &= (1+A_1A_2)X + A_1Y + \Sigma T_i \\ \xi_5 &= (1+A_1A_2) \end{aligned}\right\} \tag{15}$$

For $T_1=0.07$ second, $T_2=0.16$ second, $T_3=0.18$ second, $T_4=1.7$ seconds, $A_1=27.6$, and $A_2=3$, the ξ's shown in Table I are given. Furthermore, if these values

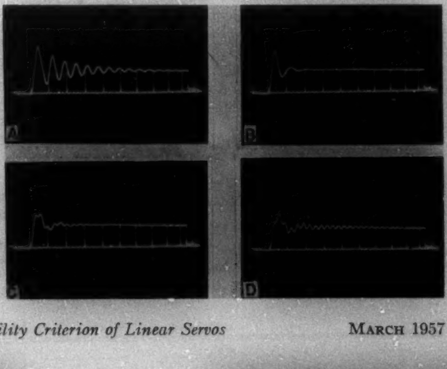


Fig. 8. Solutions of automatic voltage control system (no. 1) by repetitive analog computer (10 seconds of actual phenomena represented)

A—X=0.1 second, Y=0.03
B—X=0.1 second, Y=0.04
C—X=0.1 second, Y=0.06
D—X=0.1 second, Y=0.07

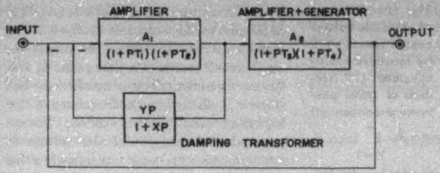


Fig. 9. Block diagram of automatic voltage control system no. 2

Table III. Coefficients of Characteristic Equation Calculated from Equation 17 (where $\xi_j = a_jX + b_jY + c_j$)

j	ξ_j	a	b	c
0	ξ_0	0.00342	0	0
1	ξ_1	0.0918	0	0.00342
2	ξ_2	0.75	7.04	0.0918
3	ξ_3	2.11	43.2	0.75
4	ξ_4	83.8	23	2.11
5	ξ_5	0	0	83.8

are substituted in equation 10, then

$$X = \frac{-46.9\gamma(2.11-0.0918\gamma) - 27.6(83.8-0.75\gamma+0.00342\gamma^2)}{\gamma\{-27.6(2.11-0.0918\gamma) + 46.9(83.8-0.75\gamma+0.00342\gamma^2)\}}$$

$$Y = \frac{(83.8-0.75+0.00342\gamma^2)^2 + \gamma(2.11-0.0918\gamma)^2}{\gamma\{-27.6(2.11-0.0918\gamma) + 46.9(83.8-0.75\gamma+0.00342\gamma^2)\}}$$

With the foregoing equations, the boundary curve can be drawn on the X-Y plane as shown in Fig. 7. The numerals near the points on the boundary curve show the values of γ. From the criterion using equation 11, the left-hand side of the boundary curve is stable facing the direction in which γ increases, but the left-hand side of this curve in the second quadrant ($X<0$) is not stable, as will be shown later. In this domain the criterion using equation 11 is not sufficient. Fig. 7 shows that if $X>0.2$, then the system is not stabilized for all the values of Y, or an upper boundary of X exists to stabilize the system. If $0<X<0.2$, there exists both upper and lower boundaries of Y for stabilization.

Figs. 8(A), (B), (C), and (D) show the initial responses of this system by a repetitive analog computer with X fixed at 0.1 second and Y increased. They correspond to points A, B, C, and D in Fig. 7. At points A and B, the system is affected by the oscillation frequency of $f_1(=\sqrt{\gamma}/2\pi)$, and at points C and D the system is affected by the oscillation frequency of $f_2(=\sqrt{\gamma_2}/2\pi)$. Table II compares the frequencies calculated from γ (i.e., $f=\sqrt{\gamma}/2\pi$) and from the results obtained by the computer (i.e., the number of waves in Fig. 8. From Fig. 8 it is seen that the closer the point on the X-Y plane approaches to boundary curve, the more the system is oscillatory, and the initial response is seriously affected by the γ near the point.

EXAMPLE: AUTOMATIC VOLTAGE CONTROL SYSTEM (No. 2)

In this section, the system shown in Fig. 9 is considered. The closed-loop transfer function of the system is given by

$$AG(p) = \frac{A_1A_2(1+Xp)}{\{(1+pT_1)(1+pT_2)(1+Xp) + A_1Yp\}(1+pT_3)(1+pT_4) + A_1A_2(1+Xp)} \tag{16}$$

A comparison of the characteristic equation (the denominator of equation 16) with equation (1) obtains $n=5$ and

$$\left.\begin{aligned}
\xi_0 &= \Pi T_i \cdot X\\
\xi_1 &= \Sigma T_iT_jT_k \cdot X + \Pi T_i\\
\xi_2 &= \Sigma T_iT_j \cdot X + T_3T_4A_1Y + \Sigma T_iT_jT_k\\
\xi_3 &= \Sigma T_i \cdot X + (T_3+T_4)A_1Y + \Sigma T_iT_j\\
\xi_4 &= (1+A_1A_2)X + A_1Y + \Sigma T_i\\
\xi_5 &= (1+A_1A_2)
\end{aligned}\right\} \tag{17}$$

For $A_1=23$ and $A_2=3.6$, and for T_1, T_2, T_3, and T_4 equal to the previous example, ξ's are calculated as shown in Table III. By use of Table III and equation 10,

$$X = \frac{-43.2\gamma(2.11-0.0918\gamma) - (23-7.04\gamma)(83.8-0.75\gamma+0.00342\gamma^2)}{\gamma\{-(2.11-0.0918\gamma)(23-7.04\gamma) + 43.2(83.8-0.75\gamma+0.00342\gamma^2)\}}$$

$$Y = \frac{(83.8-0.75\gamma+0.00342\gamma^2)^2 + \gamma(2.11-0.0918\gamma)^2}{\gamma\{-(2.11-0.0918\gamma)(23-7.04\gamma) + 43.2(83.8-0.75\gamma+0.00342\gamma^2)\}}$$

By use of the foregoing equations, boundary curves are drawn on the X-Y plane as shown in Fig. 10. For $0<\gamma<70$, the curve lies in the first and second quadrants. The stable region exists in the left-hand side of the boundary curve I, because the denominator is positive. For $70<\gamma<\infty$, the curve lies in the third quadrant, and the stable region exists in the right-hand side of the boundary curve II, because the denominator is negative. Although in these quadrants some stable regions were found by the criterion using equation 11, there exist no stable regions in the second and third quadrants by use of Hurwitz's or Hermite's theorem. That is, the stable region exists only in the first quadrant, $X>0$, $Y>0$.

Fig. 10 shows that there is a lower boundary for Y to stabilize the system, but there is no upper boundary. Thus if Y increases, the system is stabilized for any value of X. It also shows that the more X increases, the faster is the

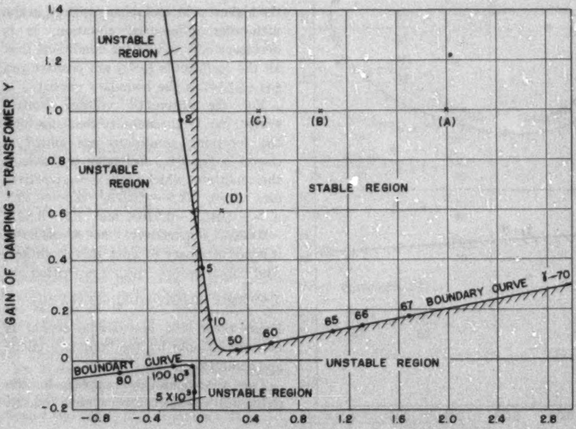


Fig. 10. Stability chart for automatic voltage control system no. 2

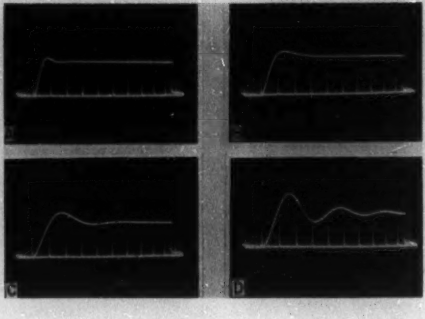


Fig. 11. Solutions of automatic voltage control system no. 2 by repetitive analog computer (10 seconds of actual phenomena represented)

A—$X=2$ second, $Y=1$
B—$X=1$ second, $Y=1$
C—$X=0.5$ second, $Y=1$
D—$X=0.2$ second, $Y=0.6$

system response, because the value of γ on the boundary curve increases. However, it is of no use to increase X indefinitely because the magnitude of γ does not exceed approximately 70($f=$ 1.33 cycles per second) at $X\rightarrow\infty$. Figs. 11(A), (B), and (C) show the initial responses of this system by the computer with Y fixed at 1 and X varied, and Fig. 11(D) shows the responses at $X=0.2$ and $Y=0.6$. They correspond to points A, B, C, and D in Fig. 10. At point C, the system is under adverse conditions, because it is affected by the frequency of the boundary curve near $\gamma=2\sim3$; but at point A, the response of the system is much improved because it approaches the curve of $\gamma=60\sim70$. At the point D, it is very oscillatory because it is very close to the curve of $\gamma=2\sim3$.

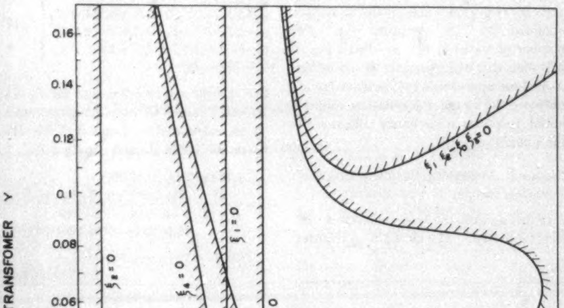


Fig. 12. Stability chart for automatic voltage control system no. 1 using Hurwitz's theorem

DISCRIMINATION OF STABLE REGIONS ON X-Y PLANE BY HURWITZ'S THEOREM

According to the foregoing example, it is clear that the criterion using equation 11 is not sufficient. Although this criterion is convenient for rough discrimination because equation 11 is calculated as the denominator of X and Y, other criteria must be used for the complete determination of the stable regions. Both Hermite's theorem and Hurwitz's theorem can be considered. As the latter is the simpler, it will be discussed here for practical use. For the former, see Appendix III.

Hurwitz's theorem requires that all the coefficients of the characteristic equation are positive as necessary conditions. In addition to this, if all the odd(even) principal minors are also positive, then the theorem is satisfied.

The highest principal minor is derived from equation 10 by eliminating γ in the right-hand side of these equations; the boundary curves are equivalent to the highest principal minor. Thus, in the fifth-order differential equation, it is necessary only to add the conditions that all the coefficients $\xi_0\sim\xi_5$ are positive and $\xi_1\xi_2-\xi_0\xi_3>0$ to the boundary curves.

For the automatic voltage control system (no. 1), the stability chart in which the foregoing conditions are added is shown in Fig. 12. Referring to Table I, the conditions which make $\xi_0\sim\xi_4$ positive, are $X>0$, $X>-0.0273$, $X>-0.122$, $X>-0.045X-0.0160$, and $Y>-3.04X-0.0765$, respectively. As $\xi_5=83.8>0$, it is not necessary to show this condition. And condition $\xi_1\xi_2-\xi_0\xi_3>0$ is written as

$$Y\lesseqgtr 0.386X+0.0527+0.00196/X(X\gtrless 0)$$

From them, the real stable region is determined only by the boundary curve and condition $\xi_0>0$ (i.e. $X>0$).

This discussion also applies to the automatic voltage control system (no. 2).

Conclusions

This paper has discussed the stability of the systems pertinent to U-V plane analysis, and it has been shown that the stable regions of such systems are very easily discriminated by Hermite's theorem. The paper has also discussed the systems whose coefficients of characteristic equation depend on two parameters, X and Y, linearly. When the optimum condition of a system is sought by use of a computer, if the stability chart peculiar to the system is known, the optimum points can be found very quickly and all stable regions accounted for.

This method can be extended, for example, when each of the two parameters in the characteristic equation is represented by γ explicitely, and the boundary curves are drawn on a plane using these two parameters as co-ordinates.

A further study of nonlinear control system is being made using this and the describing function methods.

Appendix I

Theorem: $f(z)$ is a polynomial of z having complex coefficients, and $R(z)$ and $I(z)$ are defined as shown in the following. If all the imaginary parts of the roots of equation $f(z)$ have the same sign, then both $R(z)$ and $I(z)$ have real roots and separate each other (Hermite or Biehler's theorem).

When

$$f(z)=a_0z^n+a_1n^{n-1}+\ldots+a_n$$

$$a_0=b_0+ic_0,\ a_1=b_1+ic_1,\ \ldots\ a_n=b_n+ic_n$$

$$R(z)=b_0z^n+b_1z^{n-1}\ldots+b_n$$

$$I(z)=c_0z^n+c_1z^{n-1}+\ldots+c_n$$

Using this theorem, if one compares

$f(i\omega)=u(\omega)+iv(\omega)$ with

$$f(p)=\xi_n+\xi_{n-1}p+\ldots+\xi_0p^n \tag{1}$$

then,

$$u(\omega)=\xi_n-\omega^2\xi_{n-2}+\omega^4{}_{n-4}-\ldots$$

$$v(\omega)=\omega(\xi_{n-1}-\omega^2\xi_{n-3}+\omega^4\xi_{n-5}-\ldots)$$

As $u(\omega)$ and $v(\omega)$ contain only odd or even power of ω, putting $\omega^2=\gamma$, obtains

$$\left.\begin{aligned}U(\gamma)&=\xi_n-\gamma\xi_{n-2}+\gamma^2\xi_{n-4}-\ldots\\ V(\gamma)&=\xi_{n-1}-\gamma\xi_{n-3}+\gamma^2\xi_{n-5}-\ldots\end{aligned}\right\} \tag{2}$$

Therefore, in regard to equation 2, Hermite's theorem becomes as follows: If every sign of the real parts of the roots of equation 2 is positive, then the roots $U(\gamma)=0$ and $V(\gamma)=0$ separate each other; the smallest root belongs to $U(\gamma)=0$, the number of roots of $U(\gamma)=0$ and $V(\gamma)=0$ are equal to the degree of these equations respectively, and every root is positive.

By Hermite's theorem, however, it is only indicated that the real parts of all the roots have the same sign, it is not known whether the sign is plus or minus. Thus

$$\begin{aligned}f(p)&=\xi_0(p^2+\beta_1p+\gamma_1)(p^2+\beta_2p+\gamma_2)\ldots\\ &\qquad(p^2+\beta_mp+\gamma_m),\ n=2m\\ &=\xi_0(p^2+\beta_1p+\gamma_1)\ldots\\ &\qquad(p^2+\beta_mp+\gamma_m)(p+\beta_{m+1}),\quad n=2m+1\end{aligned}$$

is put where $\gamma_j>0(j=1,\ 2,\ 3,\ \ldots,\ m)$. It is clear that if $\beta_j>0$ in this equation, then ξ_{n-2j} and $\xi_{n-(2j+1)}$ are both positive, and if $\beta_j<0$, then ξ_{n-2j} and $\xi_{n-(2j+1)}$ have opposite signs. As a result, condition 4, previously mentioned, is obtained.

Appendix II

To prove the criterion put forth under "Boundary Curves on the X-Y Plan," Tashiro's method will be used. For the characteristic equation of the stable system of $(n-2)$-degree, the following is written

$$\zeta_0p^{n-2}+\zeta_1p^{n-3}+\ldots+\zeta_{n-2} \tag{18}$$

By multiplying $(p^2+\beta p+\gamma)$ by equation 18, one can obtain the system of n-degree,

$$(p^2+\beta p+\gamma)(\zeta_0p^{n-2}+\zeta_1p^{n-3}+\ldots+\zeta_{n-2})=\xi_0p^n+\xi_1p^{n-1}+\ldots+\xi_n \tag{19}$$

which is also stable.

Then, there exist following relations between $[\zeta_0,\ \zeta_1,\ \ldots,\ \zeta_{n-3},\ \zeta_{n-2}]$ and $[\xi_0,\ \xi_1,\ \ldots,\ \xi_{n-1},\ \xi_n]$:

$$\left.\begin{aligned}\xi_0&=\zeta_0\\ \xi_1&=\zeta_1+\beta\zeta_0\\ \xi_2&=\zeta_2+\beta\zeta_1+\gamma\zeta_0\\ \xi_3&=\zeta_3+\beta\zeta_2+\gamma\zeta_1\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \xi_{n-3}&=\zeta_{n-3}+\beta\zeta_{n-4}+\gamma\zeta_{n-5}\\ \xi_{n-2}&=\zeta_{n-2}+\beta\zeta_{n-3}+\gamma\zeta_{n-4}\\ \xi_{n-1}&=\qquad\beta\zeta_{n-2}+\gamma\zeta_{n-3}\\ \xi_n&=\qquad\qquad\gamma\zeta_{n-2}\end{aligned}\right\} \tag{20}$$

If equation 20 is differentiated by β and $\beta=0$, then the following is obtained.

$$\left.\begin{aligned}\frac{\delta\xi_0}{\delta\beta}&=\frac{\delta\zeta_0}{\delta\beta}\\ \frac{\delta\xi_1}{\delta\beta}&=\frac{\delta\zeta_1}{\delta\beta}+\zeta_0\\ \frac{\delta\xi_2}{\delta\beta}&=\frac{\delta\zeta_2}{\delta\beta}+\zeta_1+\gamma\frac{\delta\zeta_0}{\delta\beta}\\ \frac{\delta\xi_3}{\delta\beta}&=\frac{\delta\zeta_3}{\delta\beta}+\zeta_2+\gamma\frac{\delta\zeta_1}{\delta\beta}\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\delta\xi_{n-3}}{\delta\beta}&=\frac{\delta\zeta_{n-3}}{\delta\beta}+\zeta_{n-4}+\gamma\frac{\delta\zeta_{n-5}}{\delta\beta}\\ \frac{\delta\xi_{n-2}}{\delta\beta}&=\frac{\delta\zeta_{n-2}}{\delta\beta}+\zeta_{n-3}+\gamma\frac{\delta\zeta_{n-4}}{\delta\beta}\\ \frac{\delta\xi_{n-1}}{\delta\beta}&=\qquad\zeta_{n-2}+\gamma\frac{\delta\zeta_{n-3}}{\delta\beta}\\ \frac{\delta\xi_n}{\delta\beta}&=\qquad\qquad\gamma\frac{\delta\xi_{n-2}}{\delta\beta}\end{aligned}\right\} \tag{21}$$

From equations 21,

$$\frac{\delta\xi_n}{\delta\beta}-\gamma\frac{\delta\xi_{n-2}}{\delta\beta}+\gamma^2\frac{\delta\xi_{n-4}}{\delta\beta}-\ldots=-\gamma(\zeta_{n-3}-\gamma\zeta_{n-5}+\gamma^2\zeta_{n-7}-) \tag{22}$$

and

$$\frac{\delta\xi_{n-1}}{\delta\beta}-\gamma\frac{\delta\xi_{n-3}}{\delta\beta}+\gamma^2\frac{\delta\xi_{n-5}}{\delta\beta}-\ldots=\zeta_{n-2}-\gamma\zeta_{n-4}+\gamma^2\zeta_{n-6}- \tag{23}$$

On the other hand, if equation 20 is differentiated by γ and $\beta=0$

$$\left.\begin{aligned}\frac{\delta\xi_0}{\delta\gamma}&=\frac{\delta\zeta_0}{\delta\gamma}\\ \frac{\delta\xi_1}{\delta\gamma}&=\frac{\delta\zeta_1}{\delta\gamma}\\ \frac{\delta\xi_2}{\delta\gamma}&=\frac{\delta\zeta_2}{\delta\gamma}+\zeta_0+\gamma\frac{\delta\zeta_0}{\delta\gamma}\\ \frac{\delta\xi_3}{\delta\gamma}&=\frac{\delta\zeta_3}{\delta\gamma}+\zeta_1+\gamma\frac{\delta\zeta_1}{\delta\gamma}\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\delta\xi_{n-2}}{\delta\gamma}&=\frac{\delta\zeta_{n-2}}{\delta\gamma}+\zeta_{n-4}+\gamma\frac{\delta\zeta_{n-4}}{\delta\gamma}\\ \frac{\delta\xi_{n-1}}{\delta\gamma}&=\qquad\zeta_{n-3}+\gamma\frac{\delta\zeta_{n-3}}{\delta\gamma}\\ \frac{\delta\xi_n}{\delta\gamma}&=\qquad\zeta_{n-2}+\gamma\frac{\delta\zeta_{n-2}}{\delta\gamma}\end{aligned}\right\} \tag{24}$$

From equations 24,

$$\frac{\delta\xi_n}{\delta\gamma}-\gamma\frac{\delta\xi_{n-2}}{\delta\gamma}+\gamma^2\frac{\delta\xi_{n-4}}{\delta\gamma}-\ldots=\zeta_{n-2}-\gamma\zeta_{n-4}+\gamma^2\zeta_{n-6}-\ldots \tag{25}$$

and

$$\frac{\delta\xi_{n-1}}{\delta\gamma}-\gamma\frac{\xi_{n-3}}{\delta\gamma}+\gamma^2\frac{\delta\xi_{n-5}}{\delta\gamma}-\ldots=\zeta_{n-3}-\gamma\zeta_{n-5}+\gamma^2\zeta_{n-7}-\ldots \tag{26}$$

From equations 22, 23, 25, and 26,

$$\frac{\delta\xi_n}{\delta\beta}-\gamma\frac{\delta\xi_{n-2}}{\delta\beta}+\gamma^2\frac{\delta\xi_{n-4}}{\delta\beta}-\ldots=-\gamma\left(\frac{\delta\xi_{n-1}}{\delta\gamma}-\gamma\frac{\delta\xi_{n-3}}{\delta\gamma}+\gamma^2\frac{\delta\xi_{n-5}}{\delta\gamma}-\ldots\right) \tag{27}$$

and

$$\frac{\delta\xi_{n-1}}{\delta\beta}-\gamma\frac{\delta\xi_{n-3}}{\delta\beta}+\gamma^2\frac{\delta\xi_{n-5}}{\delta\beta}-\ldots=\frac{\delta\xi_n}{\delta\gamma}-\gamma\frac{\delta\xi_{n-2}}{\delta\gamma}+\gamma^2\frac{\delta\xi_{n-4}}{\delta\gamma} \tag{28}$$

Substituting $\xi_j=a_jX+b_jY+c_j$ $(j=0,\ 1,\ 2,\ \ldots,\ n)$ into equations 27 and 28

$$\sum(-\gamma)^ja_{n-2j}\frac{\delta X}{\delta\beta}+\sum(-\gamma)^jb_{n-2j}\frac{\delta Y}{\delta\beta}=-\gamma\left\{\sum(-\gamma)^ja_{n-(2j+1)}\frac{\delta X}{\delta\gamma}+\sum(-\gamma)^j\cdot b_{n-(2j+1)}\frac{\delta Y}{\delta\gamma}\right\} \tag{29}$$

$$\sum(-\gamma)^ja_{n-(2j+1)}\frac{\delta X}{\delta\gamma}+\sum(-\gamma)^jb_{n-(2j+1)}\frac{\delta Y}{\delta\beta}=\sum(-\gamma)^ja_{n-2j}\frac{\delta X}{\delta\gamma}+\sum(-\gamma)^jb_{n-2j}\frac{\delta Y}{\delta\gamma} \tag{30}$$

The unit vectors in the X and Y directions are $\mathbf{i}$ and $\mathbf{j}$ respectively, and thus:

$$\mathbf{V}=\left[\left(\mathbf{i}\frac{\delta X}{\delta\gamma}+\mathbf{j}\frac{\delta Y}{\delta\gamma}\right)\left(\mathbf{i}\frac{\delta X}{\delta\beta}+\mathbf{j}\frac{\delta Y}{\delta\beta}\right)\right]_{\beta=0} \tag{31}$$

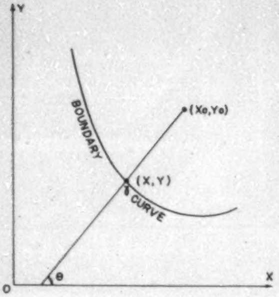


Fig. 13. Graphical determination on the X-Y plane of roots of U(γ)=0 or V(γ)=0

Therefore, if equation 31 is positive (negative), then the left- (right-) hand side of each boundary curve is stable and right- (left-) hand side is unstable facing the direction in which γ increases.

Substituting equations 29 and 30 into equation 31 obtains

$$\mathbf{V}=\frac{\delta X}{\delta\gamma}\cdot\frac{\delta Y}{\delta\beta}-\frac{\delta Y}{\delta\gamma}\cdot\frac{\delta Y}{\delta\beta}=$$

$$\frac{\left\{\sum(-\gamma)^j a_{n-2j}\frac{\delta X}{\delta\gamma}+\sum(-\gamma)^j b_{n-2j}\times\frac{\delta Y}{\delta\gamma}\right\}^2+\gamma\left\{\sum(-\gamma)^j a_{n-(2j+1)}\frac{\delta X}{\delta\gamma}+\sum(-\gamma)^j b_{n-(2j+1)}\frac{\delta Y}{\delta\gamma}\right\}^2}{\Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)}-\Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j}} \tag{32}$$

The numerator of equation 32 is always positive and so the criterion as stated under "Boundary Curves on the *X*-*Y* Plans" is obtained.

If the denominator

$$\Delta(\gamma)=\Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)}-\Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j}$$

is positive, then the left-hand side is stable, and if it is negative, then right-hand side is stable.

Appendix III

From equations $U(\gamma)$ and $V(\gamma)$, that is,

$$U(\gamma)=X\Sigma(-\gamma)^j a_{n-2j}+Y\Sigma(-\gamma)^j b_{n-2j}+\Sigma(-\gamma)^j c_{n-2j}$$
$$V(\gamma)=X\Sigma(-\gamma)^j a_{n-(2j+1)}+Y\Sigma(-\gamma)^j b_{n-(2j+1)}+\Sigma(-\gamma)^j c_{n-(2j+1)} \tag{9}$$

the following is obtained:

$$U(\gamma)\Sigma(-\gamma)^j b_{n-(2j+1)}-V(\gamma)\Sigma(-\gamma)^j b_{n-2j}=\Delta(\gamma)\left\{X-\frac{\Sigma(-\gamma)^j b_{n-2j}\Sigma(-\gamma)^j c_{n-(2j+1)}-\Sigma(-\gamma)^j b_{n-(2j+1)}\Sigma(-\gamma)^j c_{n-2j}}{\Delta(\gamma)}\right\}$$

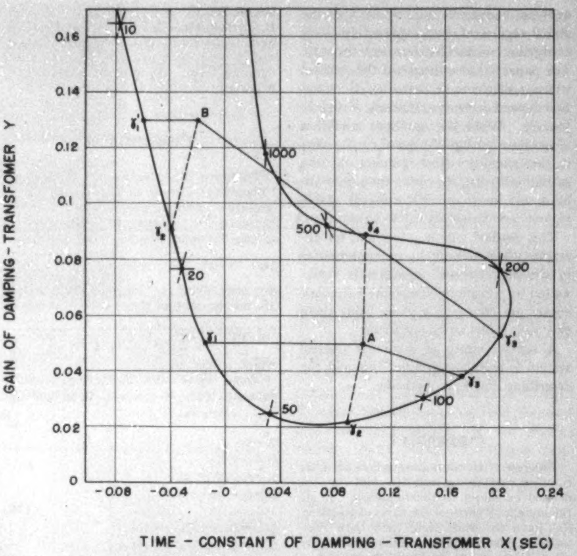


Fig. 14. Stability chart for automatic voltage control system no. 1 using Hermite's theorem

$$U(\gamma)\Sigma(-\gamma)^j a_{n-(2j+1)}-V(\gamma)\Sigma(-\gamma)^j a_{n-2j}=-\Delta(\gamma)\left\{Y-\frac{\Sigma(-\gamma)^j c_{n-2j}\Sigma(-\gamma)^2 a_{n-(2j+1)}-\Sigma(-\gamma)^j c_{n-(2j+1)}\Sigma(-\gamma)^j a_{n-2j}}{\Delta(\gamma)}\right\}$$

where

$$\Delta(\gamma)=\Sigma(-\gamma)^j a_{n-2j}\Sigma(-\gamma)^j b_{n-(2j+1)}-\Sigma(-\gamma)^j a_{n-(2j+1)}\Sigma(-\gamma)^j b_{n-2j} \tag{11}$$

By denoting any point on the *X*-*Y* plane by $(X_0,\ Y_0)$ and any point on the boundary curves by $(X,\ Y)$, then from equation 10 one gets

$$U(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j b_{n-(2j+1)}-V(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j b_{n-2j}=\Delta(\gamma)(X_0-X)$$
$$U(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j a_{n-(2j+1)}-V(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j a_{n-2j}=\Delta(\gamma)(Y-Y_0) \tag{33}$$

If $U(X,\ Y,\ \gamma)=0$, then

$$Y_0-Y=\frac{V(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j a_{n-2j}}{\Delta(\gamma)}$$

$$X_0-X=-\frac{V(X_0,\ Y_0,\ \gamma)\Sigma(-\gamma)^j b_{n-2j}}{\Delta(\gamma)}$$

and

$$\frac{Y_0-Y}{X_0-X}=-\frac{\Sigma(-\gamma)^j a_{n-2j}}{\Sigma(-\gamma)^j b_{n-2j}} \tag{34}$$

As shown in Fig. 13, when a straight line is drawn through $(X,\ Y)$ and $(X_0,\ Y_0)$, $U(X_0,\ Y_0,\ \gamma)=0$ is satisfied by γ, the right-hand side of equation 34, then is equal to the tangent of the straight line.

Similarly, γ satisfying $V(X,\ Y,\ \gamma)=0$, is given by the γ which satisfied the equation

$$\frac{Y_0-Y}{X_0-X}=-\frac{\Sigma(-\gamma)^j a_{n-(2j+1)}}{\Sigma(-\gamma)^j b_{n-(2j+1)}} \tag{35}$$

Therefore, it can be seen whether or not the values of γ obtained by this procedure satisfy the conditions 1–4 previously discussed.

This procedure as applied to system no. 1 is shown in Fig. 14. The slope of the solid line segments on the boundary curve is calculated from equation 34, and the slope of the dotted line segments is calculated from equation 35. At point *A*, the values which satisfy $U(X_0,\ Y_0,\ \gamma)=0$, are γ_1 and γ_3, and the values which satisfy $V(X_0,\ Y_0,\ \gamma)=0$ are γ_2 and γ_4. Thus *A* is a stable point. On the other hand, at point *B* the root which satisfies $V(X_0,\ Y_0,\ \gamma)=0$ is only γ_2', and so *B* is an unstable point. Therefore, the region enclosed by the *Y* axis and the boundary curve in the first quadrant is only a stable region.

System no. 2 can be considered in the same manner and it can be seen that the stable regions exist only in the first quadrant, as in the foregoing example.

Reference

1. Stabilitätsuntersuchungen in der neueren Sowjetischen Literatur, W. Hahn. *Regelungstechnik*, Munich, Germany, vol. 3, no. 9, 1955, pp. 229–31.

Vo
Pt.
19

Load and Economic Aspects of the Residential Heat Pump on Electric Utility Systems

CONSTANTINE W. BARY
FELLOW AIEE

Synopsis: This paper contains illustrations of the essential load characteristics of three types of residential heat-pump systems and their relationships to the load characteristics of the present-day full-use residential load and of the cooling load of the average home. It develops relative cost to serve the several types of heat-pump load and of the cooling load, and indicates the economic interrelationship of these loads.

AN "all-electric home," the dream and aspiration of many thoughtful men in the electrical utility industry, cannot be attained without electricity performing the job of providing comfort through heating and cooling of the home. As has been pointed out before,[1] the economic attainment of this goal is feasible through development by manufacturers of the heat pump. However, this will be a different type of load, both in its magnitude as well as in its characteristics, from the loads the electrical utility industry was acquiring in the past.[2,3] This difference will require in all phases of the utility business new approaches, new methods, and new concepts to permit supplying this new load in an economically sound manner. What is even more important, these new approaches must be instituted now in order to prevent, or at least minimize, possible future economic difficulties.

The object of this paper is to examine the basic physical and economic factors, and other considerations, which require evaluation in determining the economics of the residential heat-pump load, and to demonstrate their application to conditions of a large metropolitan system operating in the eastern part of the United States. The results obtained for that particular system will undoubtedly differ in detail from those that may be derived for other systems, but it is believed that certain general concepts underlying such evaluations are of such fundamental nature that they will be useful to others.

The load and economic aspects of the heat pump can be based on facts. With that knowledge it is then possible to exercise judgment in any consideration of pricing of the electric service to this device, which oftentimes involves the use of emotions. This paper has been prepared so as to give an opportunity to the reader to exercise his own judgment and, at the same time, to place before him the facts which cannot be dismissed as immaterial in the economic evaluation of the load of this new application in the use of electricity.

All data used in this paper reflect either the results of actual experience, or the careful research which the author and others in the industry have been conducting over many years. They reflect the fundamental concepts and well-tested techniques which have been established and have been used in studying the behavior of electric loads and the costs to serve them.[4] Actual experience has proved them to be reliable guides in determining the economics of the services rendered by systems with which the author is familiar.

The load characteristics used in this paper, for some of the types of heat pumps which so far have had little practical application, have been painstakingly calculated, in some cases hour by hour for the average coldest and warmest spells, from the actual climatological data applicable in the territory. They were then checked against the author's test experience with gas house heating[5] and air conditioning, and were further checked with data on electric house heating, air conditioning, and heat pumps, compiled by the Load Research Committee of the Association of Edison Illuminating Companies.

The load characteristics thus developed and used herein can be considered representative of an average-size home which will be significant in the economics of the heat pump in the area served by the system previously mentioned. Such a home reflects, under standard design conditions of 0 F (degrees Fahrenheit) outdoor–70 F indoor temperatures, a 55,000-Btu-per-hour heat loss; and at 98 F outdoor–75 F indoor temperatures, about half that amount in heat-dissipation requirement. It is representative of an average 6-room (3-bedroom) house, well insulated and with about 9,000 cubic feet of conditioned space.

Data on the cost-to-serve used throughout this paper embrace a method of computation involving component unit costs in an established basic cost formula of electricity supply to general classes of service.[6] Such component unit costs reflect present-day conditions, and have been derived from analyses and allocations and include all elements of system's operating expenses, taxes, allowances for annual depreciation, and an adequate return on the investment in plant required.

The subject matter in this paper covers: first, some background information on climatological data and on two cost-to-serve relationships; second, essential load characteristics of three significant types of heat-pump systems with which the author is familiar, and their relationships to the load characteristics of the present-day full-use residential load and of the cooling load of the average home; third, a development of relative cost to serve several types of heat-pump load and the cooling load, by itself; and fourth, observations concerning the question of the economic interrelationship of these loads.

Climatological Characteristics

Figs. 1(A)–1(D) depict the climatological characteristics experienced in Philadelphia, Pa. These characteristics are indicative of residential heating and cooling requirements. Basically they are the primary and the strongest factors which determine the electrical characteristics of the residential heating and cooling loads.

The concept of monthly degree-days is well known. The effective degree-hours for cooling, shown in Fig. 1(A), are defined as the weighted average of wet bulb and dry bulb temperatures, where the weight assigned to the wet

---

Paper 57-12, recommended by the AIEE Domestic and Commercial Applications Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 15, 1956; made available for printing November 1, 1956.

CONSTANTINE W. BARY is with the Philadelphia Electric Company, Philadelphia, Pa.

The author acknowledges the valuable assistance and suggestions of J. A. Thielman, W. C. Astley, E. H. Bromer, and A. Abramovitz of the Philadelphia Electric Company in the preparation of the material contained in this paper.

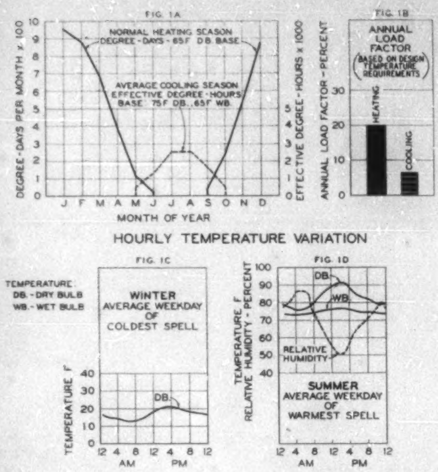


Fig. 1. Philadelphia climatological characteristics indicative of heating and cooling requirements

bulb values is two, and to the dry bulb values is one. Experience has shown that such weighted degree-hours, in Philadelphia, bears a close correlation with the electric energy rquirements for cooling of residences in that climate.

From such degree-day and effective degree-hour data and with a knowledge of design temperature requirements for the house heating and cooling installations, the so-called annual "temperature load factor" for heating is calculated at 20%, and for cooling at 6%. These values reflect optimum conditions. Under conventional methods of operation, actual load factors for these classes of electric load, based on design temperature conditions, can be expected to fall below these respective values.

While degree-days and effective degree-hours are good indicators of energy requirements for house heating and cooling, respectively, the economics of electric service to these applications are largely determined by their diversified demands over relatively short periods of time, such as for a half-hour or a one-hour interval.[6]

From many years of experience in studying the economics of house heating and cooling loads, which employ concepts of average unit cost components, the author has found that such demands are most appropriately established for conditions of the average weekday of the coldest spell for heating, and for the average weekday of the warmest spell for cooling.

The coldest spell during the heating season is defined as the period of the lowest average daily mean temperature composed of consecutive days with daily mean temperatures of 23 F or less, which, on the basis of Philadelphia Weather Bureau records, is certain to occur in that territory every year. Such average day of the coldest spell, computed from climatological data for Philadelphia for ten heating seasons from 1942–1943 to 1951–1952, inclusive, has a mean daily temperature of 17 F, a minimum hourly temperature of 13 F at 7 a.m. and a temperature of 20 F at 6 p.m. The warmest spell of the cooling season is defined as the period of highest average daily effective degree-hours of one or more consecutive days with daily effective degree-hours of 129 or more which, on the basis of Philadelphia Weather Bureau records, is a certainty of occuring in that territory every year.

The hourly summer temperatures used in this paper reflect information which was obtained as a result of the air-conditioning-load tests made by Philadelphia Electric Company in 1941 and 1947. They are the average of these two seasons, which subsequent studies have indicated to be representative of average cooling season conditions of that territory. The average day of the warmest spell, depicted in the chart, has a maximum dry bulb temperature of 91 F around 2 to 3 p.m., 83 F at 6 p.m., and a relative humidity of close to 70% at that time.

Cost-to-Serve Relationships

Fig. 2 depicts the relationships of the sum of the "demands" and "energy" components of cost of the basic cost formula[6] applicable to the residential class, and are expressed per kilowatt-hour versus annual diversified load factor. The "customer" component of cost of that formula is not included in these curves because it is not needed for this study, the purpose of which is to demonstrate the effect upon cost of the addition of the residential cooling or heat-pump loads for an existing class of residential customers.

The cost curves in this figure are expressed in per cent of the full cost per kilowatt-hour at a 45% diversified annual

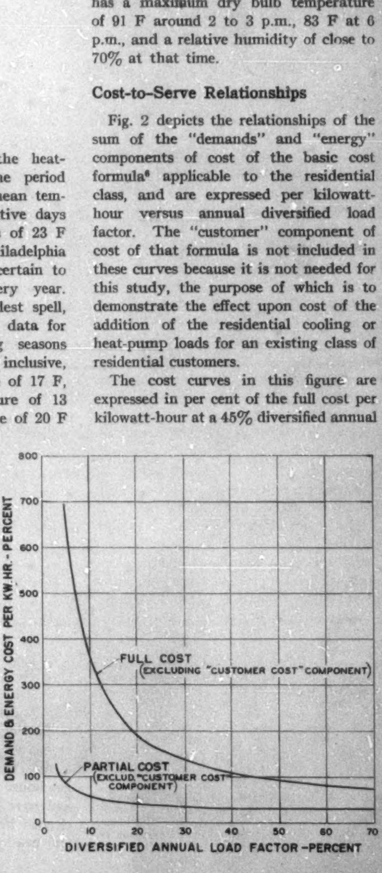


Fig. 2. Relation of demand and energy cost per kilowatt-hour and annual diversified load factor. 100% = full cost per kilowatt-hour at 45% annual load factor

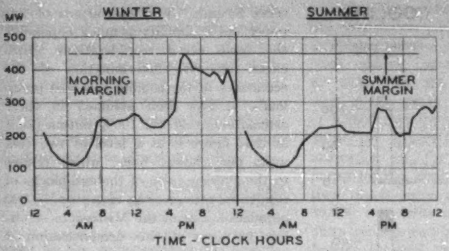


Fig. 3. Present-day residential class load for average weekday

load factor, which was selected as the 100% datum point to represent the annual load factor of the present-day total regular load of the residential class of customers, on the basis of its average weekday peak in the winter months.

The curve for full cost reflects the "customer-demand," the "class-demand," and the "energy" cost components of the basic cost formula, and thus includes all such costs from the customers' premises to and including the production system. The curve of partial cost, however, is confined to the sum of the energy and the customer-demand components of the cost formula, the latter comprising only certain cost elements of the local distribution system. Attention is directed to the hyperbolic nature of these cost relationships.

Partial cost per kilowatt-hour at 4% annual load factor, according to this curve, is equal to the full cost per kilowatt-hour at a 45% load factor. Full cost per kilowatt-hour between 25% and 30% annual load factors is from 50% to 35% higher than at a 45% load factor, and at a 5% annual load factor it is seven times the value at the 45% load factor.

Residential-Class Load Curves

Fig. 3 depicts the load curves of the residential class of service for a winter and a summer average weekday, excluding any electric heating or cooling loads for the system previously defined where people live in the summer on a one-hour daylight-saving-time basis.

The available summer margin in that illustration is about 175 megawatts without any air-conditioning load. That valley can thus be filled up with a seasonal load, such as cooling, at partial cost only. This is an important physical and economic phenomenon. It must be watched with great alertness by utility managements, because once that valley is filled up, the economic status of the cooling load changes from the partial-cost basis to the full-cost basis, and becomes in the order of seven times the partial cost per kilowatt-hour.

The available morning margin in the winter in that illustration is about 200 megawatts. This margin is available to a load, such as house heating, which peaks up in the morning when the hourly temperature is the lowest.[3] A certain amount of such load can thus be acquired at a lower demand cost responsibility than applies to its morning peak.

However, these lower costs cannot last. Once the valleys are filled up, additional load must bear the full and higher demand cost responsibilities.

Many will recall the experience of the gas industry with the gas house-heating load, which in the thirties, in many parts of the United States, appeared just to be valley filling, but in the middle forties became the peak-creating load. The consequences of the changed status of that load are too well known to need any mention here. A similar change has already happened with the residential cooling load on many electric systems and will occur on other systems in a few more years.

For the system used in this paper for illustration, the residential-class peak in the summer is still lower than in the winter.

Load Characteristics

Fig. 4 depicts the kilowatt of diversified demand, the annual kilowatt-hour of use, and the annual load factors of three types of heat pumps, of cooling, and of the so-called full-use home which comprises lighting, miscellaneous small uses, and such major appliances as the refrigerator, range, dryer, washing machine, and water heater, the latter operating under time control with a $17^1/_2$-hour charging period and completely disconnected from service during the remaining hours which constitute the

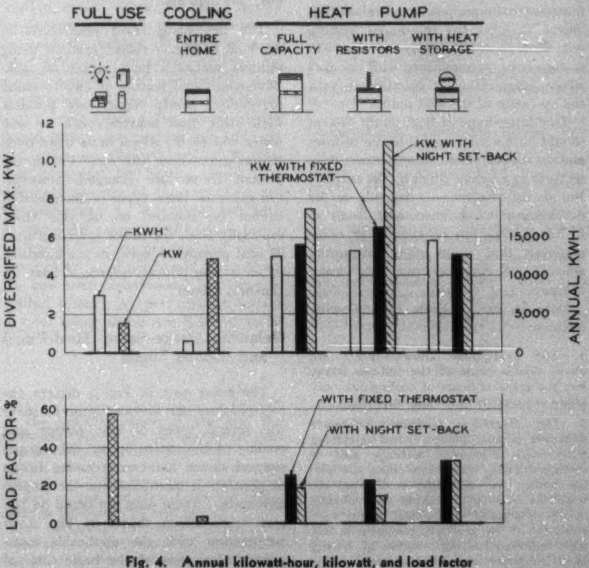


Fig. 4. Annual kilowatt-hour, kilowatt, and load factor

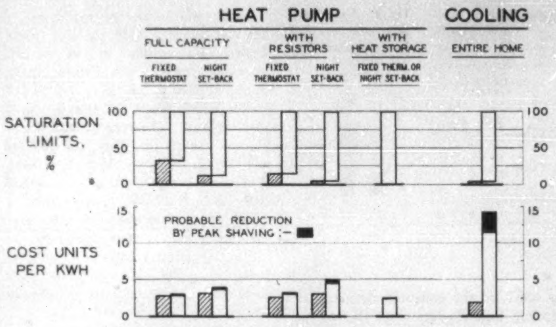


Fig. 5. **Annual cost to serve and corresponding saturation limits**

peak periods of the production and the residential-class distribution systems.

The three types of heat pumps are: (1) the conventional full-capacity heat pump, which uses air as its heat source and sink without any supplementary heat sources; (2) the conventional smaller capacity heat pump, which also uses air as its heat source and sink, and is usually sized for the cooling load and therefore requires supplementary resistors to meet the full-design heating requirements; and (3) a conventional smaller capacity heat pump as in the foregoing but equipped with heat-storage means to meet the full-design heating requirements, the supply of heat to storage being accomplished with resistors whose energization is interlocked with the operation of the heat pump.[7]

This latter type of heat-pump system should have a solid appeal to the utilities and to its customers as was more fully set forth in a former paper by the author.[1] But its full effectiveness depends on the development of suitable salts, such as hydrated disodium phosphate, or other materials that could store substantial amounts of heat in relatively small volumes.

Data in Fig. 4 show the following features:

1. The heat-pump kilowatt-hours are about double those of the full-use home, but the kilowatt-hours of cooling are only about one quarter of the full-use home.
2. The diversified maximum kilowatt demands of heat pumps, when operating with fixed thermostat settings, and of cooling are more than three times the full-use home. But attention is directed toward the pronounced effects on such demands (especially of the heat pump with resistors) of the practice, generally in use with fuel-fired heating systems, of setting back the home thermostat at night. Under that practice such electrical demands can be 80% higher than when such heat pumps operate with fixed thermostat settings day and night.
3. The full-use home provides close to a 60% diversified annual load factor. Such load factors for the heat pumps, operating with fixed thermostat settings, vary from 23% for heat pumps with resistors to 33% for heat pumps with heat storage. But when these heat pumps operate with night setback, these load factors deteriorate to levels below 20%, except for the heat pump with heat storage which remains at 33%. The diversified annual load factor of cooling is extremely low; less than 5%.

The heat-pump load characteristics depicted in Fig. 4 reflect systems using outdoor air as a heat source or sink. Diversified load factors of two to three percentage points higher are possible with other heat sources, such as well water and earth, which have more even temperatures than the air. Unless the present trends are changed, however, too many of these types of installations cannot be counted on at this time, especially since the major manufacturers of heat pumps are going in the direction of air as the primary source of heat for the heat pump.

Relative Cost-to-Serve Heat-Pump and Cooling Loads

The lower part of Fig. 5 depicts the relative cost per kilowatt-hour to serve the several types of heat pumps and cooling of the entire home. The upper portion shows the corresponding limits of permissible saturation for the respective costs. These data are based on the cost relationships shown in Fig. 2 in conjunction with the applicable component unit costs of the basic cost-to-serve formula. For the purpose of this paper and to simplify as much as possible the basic broad considerations which enter into the determination of the economics of the cooling and heat-pump loads, it has been assumed in these calculations that the 100% datum point in that figure is at a level of two cost units per kilowatt-hour, which relates to the average price of the end-blocks of residential rates in cities with over 500,000 population east of the Mississippi River. For purposes of this demonstration it is assumed that this level of cost units is fully compensatory for all additional uncontrolled energy used by residential customers which provide an annual diversified load factor of 45%. It should be realized, of course, that actual values will vary from system to system, and those used herein are illustrative only.

On the basis of these assumptions, the following will be seen from the data depicted in this figure:

1. The lowest relative cost per kilowatt-hour obtains for the heat pump with heat storage, and this cost applies over the entire range of possible saturations, from 0 to 100%, and is independent of customers' practices in regard to home thermostat settings day and night. This relative cost is in the order of $2^1/_2$ cost units per kilowatt-hour.
2. For short-term considerations, during the morning valley-filling period, the relative cost per kilowatt-hour for the other types of heat pumps is in the order of $2^3/_4$ cost units per kilowatt-hour under fixed thermostat settings. However, this condition prevails only for a limited range of heat-pump saturation. Under night setback of thermostats, the relative cost per kilowatt-hour of the full-capacity heat pump and of the type with supplementary resistors goes over the three cost unit mark; and the saturation limits are reduced to one third of the values possible with fixed thermostat settings.
3. For long-term considerations, beyond the first limited-saturation range, relative costs per kilowatt-hour of heat pumps (other than with heat storage) hit close to the $3^1/_4$ cost unit level with fixed thermostat settings; and reach the four and five cost unit level with night setback.
4. The relative cost per kilowatt-hour for cooling only during the valley-filling period is a little below two cost units per kilowatt-hour for the first $4^1/_2$% saturation of fully cooled homes, or their equivalent in room coolers. However, after the summer valley is filled with that relatively small percentage saturation, the relative cost to serve the cooling load alone is up to a level around 14 cost units per kilowatt-hour. And this is where the economic significance of the heat pump becomes apparent.

Before proceeding, attention is directed to some additional features depicted in Fig. 5. All relative cost values hereinbefore mentioned applied to the present-day or conventional methods of service supply. By use of unconventional methods, such as through peak shaving, for example, with either diesel engines or gas turbines connected to primary buses at substations, preliminary studies to illustrate their effects show that:

1. For some years to come, when load valleys are being filled for the indicated saturation limits of heat pumps and cooling, no saving is probable in their cost to serve by the unconventional as compared to the conventional method of electricity supply.

2. Beyond the valley-filling period, there are indications for probable savings, which are relatively small, in the order of one tenth of a cost unit per kilowatt-hour, for heat pumps operating under fixed thermostat settings, increasing to five-tenths of a cost unit per kilowatt-hour for heat pumps with resistors when they operate with night setback. For cooling alone the reduction could be as high as three cost units per kilowatt-hour.

3. However, it is to be noted that even with this type of peak shaving, the minimum relative cost to serve the heat-pump load, which is obtainable with heat storage, is in the order of $2^{1}/_{3}$ cost units per kilowatt-hour. For the other types, the minimum, under the long-term considerations, is in the order of three cost units per kilowatt-hour with fixed thermostat settings, and rises considerably above $3^{1}/_{2}$ cost units per kilowatt-hour with night setback.

4. On the basis of the foregoing illustrations, it will be seen that even for the immediate conditions and for the indicated limits of saturation, the relative cost to serve the average type of heat-pump load can be expected to be at a level of around $2^{3}/_{4}$ cost units per kilowatt-hour.

Interrelationship of Cost-to-Serve Heat-Pump and Cooling Loads

To complete the analysis, the over-all economic aspect of the interrelationship of the costs of electric service for cooling only and for the combined function of cooling and heating by means of the heat pump, will now be investigated. Based on established energy-use data and the illustrations used for the relative cost levels under the conventional method of service supply, the ultimate total relative cost-to-serve for cooling, by itself, for the average-size house used in this paper will be in the order of 23,800 cost units a year, or 14 cost units per kilowatt-hour. For limited saturation of the average-type heat pump for the same-size home, the relative cost-to-serve can be expected at 35,200 cost units a year, or $2^{3}/_{4}$ cost units per kilowatt-hour. Thus, the difference between these relative costs of 11,400 cost units divided by the difference in the corresponding energy uses, results in an annual increment cost per kilowatt-hour for heating only, over that for cooling, of one cost unit per kilowatt-hour. These are the economic facts.

From here on, it becomes a matter of business judgment on how to split these costs-to-serve between the cooling and the heating functions as a basis for establishing pricings for each of these functions to provide the optimum results for acquiring these loads. From practical considerations, the price of electricity used by the heat pump for heating need not be as low as indicated by the afore-indicated increment cost. At the same time the price for cooling, from many considerations, could not be as high as indicated by the afore-shown relative costs. Since the heat pump for heating uses the capacity required for cooling, perhaps an equal split of the "demand" component of costs for cooling between itself and heating may be feasible. Under these conditions the relative pricing for home cooling would reflect an average of, say, six cost units per kilowatt-hour. The relative increment for heating could then reflect an average of $2^{1}/_{4}$ cost units per kilowatt-hour. At this level, and with a heat-pump-system performance factor as low as $2^{1}/_{3}$, the heating cost will be much closer to, although still higher than, the cost with conventional fuels at prevailing prices. At performance factors of over three, which are believed to be attainable, it would be within striking distance of being a standoff.

The levels of the equivalent average values per kilowatt-hour just mentioned, can be accomplished most effectively through some form of demand-energy type of rate structure for the residence service, notwithstanding the many difficulties and disadvantages that are inherent in such a rate structure when applied in a mass market such as that composed of the hundreds of thousands of residential customers of a given system. It is here that new approaches, new methods, and new concepts are needed to solve successfully the economic implications of these loads.

There are indications that such new approaches will be forthcoming. But concomitantly, there must be further intensification in research and development of the device itself for improvements in its performance factor and annual load factor. Both are within the province of the manufacturing industry. Improvements in the annual load factor are believed to be possible of accomplishment only through the use of some form of heat storage,[1] the practical significance of which is dependent upon successful development of materials providing economical and dependable means for such storage.

References

1. RESIDENTIAL HEAT PUMP EXPERIMENTS IN PHILADELPHIA—SUGGESTED POSSIBILITIES FOR PRACTICAL APPLICATIONS, **Constantine Bary.** *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 361-65.

2. THE HEAT PUMP—ITS SIGNIFICANCE AS A POTENTIAL ELECTRIC LOAD, **Constantine Bary.** *Ibid.*, vol. 68, pt. I, 1949, pp. 1–8.

3. THE HEAT PUMP LOAD—GOOD OR BAD? **Constantine Bary.** *Refrigerating Engineer*, New York, N. Y., May 1950.

4. LOAD-STRUCTURE OF A MODERN ELECTRIC UTILITY SYSTEM, **Constantine W. Bary.** *Proceedings*, American Power Conference, Chicago, Ill., vol. XVI, 1954.

5. CHARACTERISTICS OF GAS HOUSE HEATING LOAD, **Constantine Bary.** *American Gas Association Monthly*, New York, N. Y., Mar. 1949.

6. LOAD CHARACTERISTICS OF "SMALL LIGHT AND POWER" CUSTOMERS—THEIR ECONOMIC SIGNIFICANCE, **Constantine W. Bary.** *Electric Light and Power*, Chicago, Ill., Oct. 15, 1955.

7. HEAT PUMP APPARATUS, **Constantine Bary.** *U. S. Patent no. 2,723,083*, Nov. 8, 1955.

◆

Discussion

P. F. O'Neill (General Electric Company, Bloomfield, N. J.). The development of a large residential heat-pump market in any area or region of the United States is not dependent upon just one economic factor. Rather, the rate of growth is influenced by the combined effect of at least three economic considerations.

Mr. Bary's excellent paper covers one of the important analyses, the load and economic aspects of the residential heat pump on electrical utility systems.

The other two primary economic considerations involve, first, the manufacturers of heat pumps, and second, the end customers or purchasers of heat pumps.

In discussing Mr. Bary's paper it is important to briefly point out some of the factors that must be examined by the manufacturers and then those pertaining to the end customers.

To obtain a mass residential heat-pump market, manufacturers must produce a product that can be competitively priced with other year-round air-conditioning systems. It must conform to or surpass the sales features, application flexibility, and functional benefits offered by other similar systems. The heat pump must meet the space requirements which the home owner will tolerate and its annual operating costs must be reasonably comparable, if not entirely comparable, to other heating and cooling systems.

The third economic consideration deals with home owners. Within any marketing area, home owners must want and must purchase year-round air-conditioning equipment; that is, they must purchase and install heating and cooling at the same time.

In addition to the residential market there are other markets where dollars and cents influence the growth of the heat-pump business. Here again economic factors determine the size of the commercial and industrial heat-pump markets.

Mr. Bary's paper presents an excellent analysis of the residential heat pump on the electrical utility system. He has pointed out that the residential heat-pump load is a different type of load and that its economics become even more important when a utility system operates under summer system peak conditions.

The phrase "operates under summer system peak conditions" is a very important one. It cannot be omitted in any analysis.

Mr. Bary's paper also emphasizes that the economics of such a load cannot be based on stagnate conditions. As he has shown, it requires long-range thinking and of long-range information.

To make such an analysis it is essential that the examiner have yearly sales forecasts of the air-conditioning industry and that this information is adapted to the area being studied.

All of us probably have read in our daily newspapers and in our trade magazines the annual sales forecasts of the air-conditioning industry, but we have seldom encountered a cumulative sales forecast. It is the cumulative sales forecasts in which utilities are interested and such a forecast is essential in developing an economic study.

Present industry forecasts indicate that in the next 10 years (1957 through 1966) room coolers and packaged air-conditioning unit sales will produce a cumulative connected load of approximately 95,000,000 kw. The major portion of this 95,000,000-kw figure is applicable to residential load growth, and a considerable portion of this total 95,000,000-kw connected load will appear in the eastern areas.

Practical saturation figures, with respect to the air-conditioning load, must be determined for a utility's service area by the people making the economic study. With these data and a knowledge of the growth of the heat-pump market and other residential loads, the resulting study should produce a practical solution.

Mr. Bary has stated that "it becomes a matter of business judgment on how to split these costs to serve between the cooling and heating functions as a basis for establishing pricings for each of these functions to provide optimum results for acquiring these loads."

The term "business judgment" can be associated with a number of topics, one of which covers the pricing of the product. In his paper Mr. Bary has stated that it can influence the unit cost per kilowatt-hour. Consequently readers might ask themselves these three questions. The answers require business judgment.

1. Should a home owner using an electric heating and cooling system pay the same unit cost per kilowatt-hour for his heating and cooling usage as his neighbor pays per kilowatt-hour for electric cooling, even though he uses another type of heating?

For example: Should both customers pay 2¢ per kilowatt-hour?

2. Should the all-electric heating and cooling customer pay a lower unit cost per kilowatt-hour for both his heating and cooling usage than his neighbor pays per unit usage for electric cooling when using a nonelectric heating system?

For example: Should the all-electric customer pay 1.5¢ and the other customer pay 2¾¢ per kilowatt-hour?

3. Should all home owners operate with the same unit cost for cooling usage while those using an electric heating system operate with a lower unit cost during the heating season?

For example: Should both customers pay 2¾¢ per kilowatt-hour for cooling and the electric heating customer purchase his heating usage at 1.5¢ per kilowatt-hour?

To some people, one or more of these questions may sound like new ideas. They are not new. They are questions the answers to which involve business judgment as well as an economic analysis. Some utilities have already exercised business judgment with respect to these questions.

Judgment factors are not just applicable to the application of the economic analysis developed for a utility system or utility systems. Business judgment must also be exercised by all the people connected with the heat-pump business and these judgment factors must all be weighed together when a person examines the heat-pump market.

We can talk about and discuss utility economics with respect to all types of heat-pump systems, such as heat storage or possibly special refrigerants. But when we do this we cannot exclude the cost to manufacturer and market such systems. The omission of an important control factor could show a market which does not truthfully exist.

The size of the residential heat-pump market in any region or, as a matter of fact, in any area depends upon a number of economic factors and the usage of good business judgment not in just one field but in the combined fields of manufacturing, marketing, and utility operations.

The public wants and will buy cooling without reference to the unit cost per kilowatt-hour.

On the other hand, the purchase of heating is strictly competitive. The all-electric home is solely contingent today upon the competitive heating cost. This is the area where business judgment in establishing unit costs per kilowatt-hour must be realistic.

Constantine W. Bary: It is gratifying to learn from Mr. O'Neill's discussion of my paper that he realizes the interdependence of the economics of utilities, manufacturers, and customers upon the growth possibilities of the residential heat-pump load. However, from what I have seen to date, the direction taken by major manufacturers in the development of the heat pump has been anything but conducive to the improvement of its load characteristics on the electric supply systems and to an increase in its over-all efficiency of performance toward levels of theoretical feasibility. Both of these features are essential parts for a wide acceptance by the public of this device.

As of this time, few, if any, improvements are noticeable in the performance characteristics of residential heat-pump systems as now marketed; and in the field of their annual load factors, the introduction of the relatively large amounts of supplementary resistance heaters, especially in heat pumps using air as a heat source, has tended to lower an already inherently low annual load factor of the device. Obviously, this becomes a further deterrent toward obtaining a low unit cost of electricity supply to this device as will be seen from Fig. 2 of the paper. Improvements in both of these features are within the province of the manufacturing industry; but it is becoming quite apparent that they can be achieved only through an intensification of research on certain fundamental aspects, such as heat storage, and through further developments in engineering and manufacturing of component parts.

The time occurrence of utilities' service peaks, referred to by Mr. O'Neill, whether in the winter or in the summer, is an important consideration in the utility systems' economics of the heat-pump load. Its significance, however, exists only for a relatively small degree of saturation of this device because the magnitude of the individual kilowatt load of this device, after reflecting diversity, is many times that of the diversified maximum kilowatt load of the present-day use, or even of a "full-use" home, as defined in the paper. Thus, from the standpoint of long-range thinking, mentioned by Mr. O'Neill, the economics of electricity supply to the heat-pump load will be determined largely by its own load characteristics, as depicted in the paper; and this is a cardinal part in any 3-way economic considerations he advocates.

In regard to Mr. O'Neill's three suggested questions, as has been pointed out in the paper, with the wide variation possible in the relative cost to serve various types of heat pumps, of cooling, and of other loads in the residential field, together with the wide variety of operating conditions that may be encountered in the characteristics of these loads, I believe that, notwithstanding the difficulties present, a proper solution rests in the establishment of some form of "capacity-energy" type of rate structure for the residential service for customers having more than a predetermined amount of use, either kilowatt-hour-wise, or kilowatt-wise, or both. To reflect automatically the relative significance of diversity between various types of residential loads, I proposed not long ago[1] that the kilowatt demand determination in such a rate structure be based on a long time-interval of registration, in the order of 4 hours, in place of the 15-minute, 30-minute, or even 60-minute intervals now commonly in use.

REFERENCE

1. A TEN-YEAR RESIDENTIAL LOAD-BUILDING PROGRAM, **Edwin Vennard.** *Electric Light and Power*, Chicago, Ill., June 15, 1956.

A 2-Dimensional Feedback Control System

PHILIP SARACHIK
STUDENT MEMBER AIEE

J. R. RAGAZZINI
MEMBER AIEE

THERE is a class of control systems whose objective is to control precisely and maintain a relationship between several dimensions. A simple example of this type of system is the contour-cutting lathe-tool control where the cutting edge of the tool is made to trace out a path in two dimensions, one in the direction of the bed movement and the other in the direction of the tool feed. One approach to the synthesis of a system to control the movement of the tool is to cause the bed to progress either at constant speed or at a programmed speed consistent with the particular cutting operation being performed. In any case, the movement in one dimension of the tool movement is open-cycle and the burden of following the prescribed path is placed entirely on the system which controls the depth movement of the tool. The latter control is usually closed-cycle and the system must have dynamic response to maintain the tool position to within given specifications as the bed is moved in its programmed path.

Other problems of similar nature occur in the guidance of a vehicle along a chosen 2-dimensional track. In this case it is desired to maintain the vehicle on the prescribed track by steering and thereby changing the heading. Extensions to 3-dimensional space paths can also be envisioned.

The common property of systems of this type is the fact that for each and every value of one of the dimensions, say x, the other variable, say y, must assume a prescribed value within some practical specification. It is important to note that time does not appear explicitly in the requirement and is considered only as a parameter. What is important is the maintenance of the relationship between x and y at all times. Naturally, in the interest of getting the job done, it is desired to complete the tracing out of the space path in a minimum time consistent with other limitations such as maximum speeds, torques, or other limitations.

A procedure is suggested here which would relate the variations of the two variables in accordance with some prescribed logic. This can best be explained by reference to Fig. 1 in which a desired path in x and y is to be traced out. The relationship between x and y can be stored on such devices as cams, digital memory devices, or possibly as a mathematical function. In any case, the storage device yields the desired value of y at its output when an input x is applied. As stated before, the object of the system is to cause the controlled element to trace out this path. As in all control systems, there will be dynamical errors due to transients or steady-state following errors as shown in Fig. 1 by the difference between the solid or desired path and the actual or dashed path.

In synthesizing the system it is assumed that a specification can be set up for the maximum allowable value of e_y. In addition, there may be other significant limits such as the maximum value of the velocity in the x- and y-directions. It will be noted here that no specification is given as to when the controlled element should be at any given point in the path, but only that it should be on or near the path within a specified amount. By eliminating time as a primary variable, it will be noted that the requirement placed on the controlled element is one of meeting the functional relationship between the two dimensions under consideration.

The logic which will be applied to this problem is that of cross-coupling the x- and y-control systems in such a fashion that the error in y affects both the x- and y-control loops. This cross coupling can be relatively simple in the case of the contour-tool controller. For instance, the velocity in the x-direction can be reduced proportionally to the magnitude of the error, e_y until it reaches zero when the error reaches its maximum allowable value. As will be shown later complete standstill is never reached in the steady state although during transient conditions it is possible.

The material which follows indicates a technique for synthesizing a control system which fulfills the foregoing objectives. More important, the system should be capable of accepting any functional relationship, thus eliminating the need for programming or adjusting the system as the desired function is changed.

Description of System

The system which will be discussed here is shown schematically in Fig. 2. The desired function to be followed by the controlled elements is stored in a data store of any form. When the store is interrogated by a signal x_c, it will deliver a signal y_R which is related by the desired function $f(x_c)$. The error detector compares the desired value of y with the actual value y_t and a control error e_y is generated by taking the difference. The interaction between the two dimensions is provided by a command distributor, D_x and D_y, which apportions the command to the x- and y-controlled elements G_x and G_y. It is noted that the input e_2 to the controlled element G_y is limited to a maximum value of L by D_y. In this way, an upper speed of movement in the y-direction can be set in at will. As will be seen later in an example, the input V sets the maximum value of the input to the controlled element G_x. In this case, the maximum speed is reached when the control error e_y is zero.

The exact operation of a system like this is dependent on the relationships which are set up in the command distributor. A very useful command distribution function is shown in Fig. 3 where the combined effects of the command distributor and limiters are shown. In the x-channel, the command to the controlled element G_x is at its maximum when the control error e_y is zero. This means that the speed in the x-direction is at its maximum value. As the control error e_y becomes larger, the magnitude of the command in the x-channel is decreased until it reaches zero when e_y reaches its maximum allowable value. This means that the speed in the x-direction is zero and the tool carriage stands stationary until the error has been reduced to below its maximum allowable value.

In the y-channel, the command is linearly related to the control error as in any linear control system except that it is limited to some desired maximum

---

Paper 57-184, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 4, 1956; made available for printing December 6, 1956.

PHILIP SARACHIK is a student and J. R. RAGAZZINI is a professor at Columbia University, New York, N. Y.

This research was supported by the United States Air Force under Contract no. *AF 18(600)677*, monitored by the Office of Scientific Research, Air Research and Development Command.

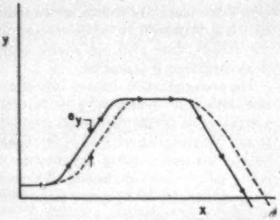


Fig. 1. Solid curve gives desired spatial path; dotted curve gives actual spatial path

value L. A superficial examination of the system indicates that the y-loop is relatively straightforward and should operate in the same manner as a conventional system for incremental changes. On the other hand, in view of its interaction with the x-channel, predictions as to its performance are not readily made. The transmission through the command distributor D_x is nonlinear since the transmission is dependent on only the magnitude of the control error. In addition, the x-loop contains the function $f(x_c)$ as one of its elements which, for the general case, is nonlinear. Therefore a complete and thorough analysis of the system for the general case where the function is not explicit is extremely difficult and probably not amenable to ordinary methods. On the other hand, considerable light can be shed on the performance of this system by studying it from two points of view. The first viewpoint looks at the over-all nonlinear system to determine its spatial performance, which is of primary interest. This approach disregards time and so dynamic performance cannot be evaluated. It does, however, give a clear picture of the actual x-y curve which the system traces and its relation to the desired curve $f(x_c)$. The second approach involves the study of an incremental linear model of the same system, and will show the system's dynamic behavior in a linear region of operation. The conclusions from such investigations would have to be verified by means of computer studies case by case.

Phase-Plane Analysis of System

For purposes of discussion, it will be assumed that the system is that shown in Fig. 2 and that the command distributer has a characteristic shown in Fig. 3. The controlled elements G_x and G_y are assumed to be simple integrators. The set of equations which relate the variables are

$$y_R = f(x_c) \tag{1}$$

$$e_1 = V - D_x(e_y) \tag{2}$$

$$e_2 = \dot{y}_c = D_y(e_y) \tag{3}$$

$$\dot{x}_c = K_x e_1 \tag{4}$$

$$e_y = y_R - y_c \tag{5}$$

The solution of the foregoing equations for x_c and y_c for any given function $f(x_c)$ describes the system performance. However, the thing of interest in this system is how y_c relates to x_c rather than how either varies in time. The representation of x_c and y_c in the phase plane readily gives this type of information. Writing the expressions for $\dot{x}_c$ and $\dot{y}_c$ in terms of functions of x_c and y_c gives

$$\dot{x}_c = K_x\{V - D_x[f(x_c) - y_c]\} \tag{6}$$

$$\dot{y}_c = D_y[f(x_c) - y_c] \tag{7}$$

which in turn gives the expression for the slopes of trajectories in the phase plane as

$$\frac{dy_c}{dx_c} = \frac{D_y[f(x_c) - y_c]}{K_x\{V - D_x[f(x_c) - y_c]\}} \tag{8}$$

Now inserting the appropriate expressions for D_x and D_y for each of the four linear regions indicated in Fig. 3, the results for trajectory slopes in each region of the phase plane are

$$\text{Region 1:} \quad \frac{dy_c}{dx_c} = -\frac{L}{0} = -\infty \tag{9}$$

$$\text{Region 2:} \quad \frac{dy_c}{dx_c} = \frac{K_y[f(x_c) - y_c]}{K_x\{V + M[f(x_c) - y_c]\}} \tag{10}$$

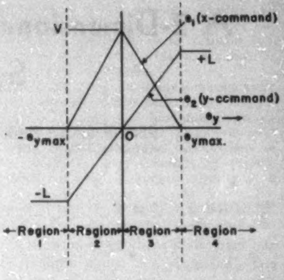


Fig. 3. Command distributor characteristic

$$\text{Region 3:} \quad \frac{dy_c}{dx_c} = \frac{K_y[f(x_c) - y_c]}{K_x\{V - M[f(x_c) - y_c]\}} \tag{11}$$

$$\text{Region 4:} \quad \frac{dy_c}{dx_c} = \frac{L}{0} = \infty \tag{12}$$

where K_y is the slope of D_y and M the magnitude of the slope of D_x in regions 2 and 3. It is noted immediately from equations 9 and 12 that if due to initial conditions or disturbances the system finds itself operating in regions 1 or 4, the system will move directly toward the curve $f(x_c)$ without advancing in x_c until the error e_y falls below $e_{ymax} = V/M$. Equations 10 and 11 cannot easily be integrated to give explicit expressions for the trajectories in regions 2 and 3. They do however completely determine these trajectories, which may be sketched quite easily. It is further noted that this may be done for any arbitrary single-valued function $f(x_c)$ because of the simple way it appears in equations 10 and 11.

A sketch of the phase plane for the $f(x_c)$ of Fig. 4 is shown in Fig. 5. Here it can be seen that no matter at what point in the phase plane the system starts, it will move toward the desired curve $f(x_c)$ and then follow it within the maximum error specification.

This particular example was done on

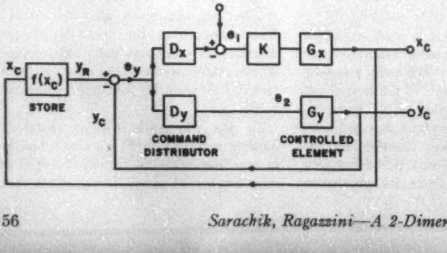


Fig. 2. Block diagram of typical 2-dimensional control system

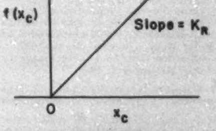


Fig. 4. Desired function $f(x_c)$ for model under discussion

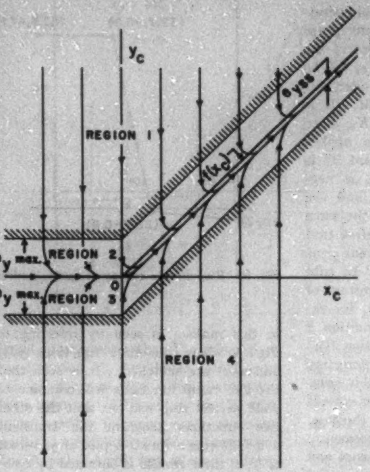


Fig. 5. Phase-plane trajectories for $f(x_c)$ in Fig. 4

the assumption that both e_1 and e_2 start limiting at $e_y = e_{y\max}$; actually this need not be the case. In practice if e_1 in Fig. 3 does not limit exactly at zero, the denominator of equations 9 and 12 will not be zero and dy_c/dx_c will not be infinite but will be equal to some large finite value in regions 1 and 4. This, however, obviously does not change the system performance drastically. Similarly the value of e_y at which e_2 limits is relatively unimportant. If in Fig. 3 e_2 limits for $e_y < e_{y\max}$ there will be only a slight change in the trajectories at the outer edges of regions 2 and 3, where as if e_2 limits for $e_y > e_{y\max}$ there will be no change at all.

The actual steady-state following error can be obtained for any region in which $f(x_c)$ is linear. In such a region of the phase plane $f(x_c)$ has a constant slope and all trajectories will approach a path which when followed maintains the error constant. In Fig. 5 this is the path with slope equal to K_R. Therefore substituting K_R for dy_c/dx_c in equation 11, the actual following error is found to be

$$e_{yss} = f(x_c) - y_c = \frac{K_R K_x V}{K_y + K_R K_x M} \tag{13}$$

It is apparent from this result that the following error can never exceed $e_{y\max} = V/M$ and that the controlling factor in reducing this error is the y loop gain K_y, as might be expected. It is noted that for the case where e_2 limits before e_1, it is possible that the steady-state trajectory does not lie in region 3 as previously defined. It might instead lie in the portion of region 3 where e_2 is limited. In this case the numerator of equation 11 is replaced by L and the steady-state following error becomes

$$e_{yss} = e_{y\max} - \frac{L}{K_x K_R M} \tag{14}$$

Analysis of Linear Incremental Model

In order to linearize this system, it is first necessary to consider the function relating y_R to x_c as piecewise linear in the regions under consideration. The functional relationship of Fig. 4 is therefore used for this analysis and in the region of interest ($x_c > 0$) is given simply by

$$y_R = K_R x_c \tag{15}$$

Similarly, because y_R and x_c are linearly related, the Laplace transform of these two variables is

$$Y_R(s) = K_R X_c(s) \tag{16}$$

The next relation to be considered is that of the D_x command distributor. The distribution function shown in Fig. 3 is given by the relation

$$e_1 = V - M|e_y| \tag{17}$$

It is noted that only the magnitude of the control error, e_y, is used. Since now only operation in a completely linear region is being considered, it is assumed that the error does not change sign during the transient operation of the system. This assumption is valid, since this analysis further assumes that the system is started at x_c equal to zero and is suddenly turned on by applying the signal V. This initiates the x translation. Under these conditions then, the expression for e_1 can be given by

$$e_1 = V \mp M e_y \tag{18}$$

where the upper sign applies if the error is positive and the lower one if negative. The Laplace transform therefore is given by

$$E_1(s) = V(s) \mp M E_y(s) \tag{19}$$

where it is implicit that V is applied when the system is turned on at zero time. The constant M is as before and is so chosen that the command e_1 is zero when the error e_y reaches its maximum allowable magnitude.

Again, since the incremental analysis is concerned only with operation for small errors, the effect of the limiter is neglected and the y-channel command distributer is linear. Thus

$$E_2(s) = K_y E_y(s) \tag{20}$$

The relations for the output variables, for the case in which no limiting occurs in terms of the transfer function of the controlled element, are given by

$$X_c(s) = K_x G_x(s) E_1(s) = K_x G_x(s)[V(s) \mp M E_y(s)] \tag{21}$$

$$Y_c(s) = K_y G_y(s) E_y(s) \tag{22}$$

Again assuming the controlled elements G_x and G_y, to be simple integrators with transfer function $1/s$, when equations 16, 21, 22, and the Laplace transform of equation 5 are combined, there results

$$X_c(s) = \frac{K_x(s + K_y)}{s[s + K_y \pm K_x K_R M]} V(s) \tag{23}$$

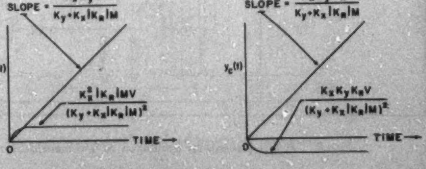


Fig. 6. x_c and y_c time response for linear incremental model

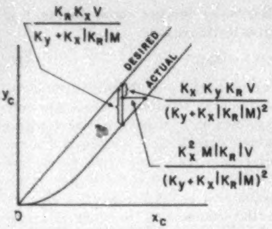


Fig. 7. Spatial response of linear model

$$Y_c(s)=\frac{K_x K_y K_R}{s[s+K_y \pm K_x K_R M]}V(s) \qquad (24)$$

where again the upper sign applies if the error e_y is positive and the lower sign if negative. Two things are noted here: First, the dynamic coefficients of the system are functions of the slope K_x of the reference function $f(x_c)$. Thus to retain the same dynamic performance for different slopes, it would seem to be desirable to make the quantity $K_x K_R M$ invariant. Since this would require a variable gain control on K_x in practice it is simpler to just choose the y-loop gain K_y, to be large enough in comparison with the product $K_x K_R M$ for the variations in this quantity to have only a negligible effect on the system dynamics.

The second point to be noted is that the dynamic coefficient $(K_y \pm K_x K_R M)$ actually has the same value for either positive or negative steady-state error. If equations 5, 16, 23, and 24 are solved for $E_y(s)$, the result is

$$E_y(s)=\frac{K_R K_x}{s+K_y \pm K_x K_R M}V(s) \qquad (25)$$

It can be seen from this that for positive K_x and $V(s)$, which is the normal operating condition, in the steady state e_y will have the sign of K_R. Therefore in the steady state for positive slopes e_y will be positive and conversely. Consider now the constant $T=K_y \pm K_x K_R M$. It is easily seen that for a positive K_R, since e_y is positive, the upper sign applies and the value of the constant T is $K_y + K_x K_R M$. If however, K_R is negative then e_y is also negative and the lower sign applies, resulting in the same value for T. It is evident therefore that only when e_y and K_R have different signs will the term $K_x K_R M$ actually be subtracted from K_y in the time constant of the system. This would occur, for example, during the transient condition if the slope of the reference function $f(x_c)$ suddenly changes sign. An interesting operating condition exists when this situation develops if the $K_x K_R M$ term exceeds the value of K_y, as may well be the case. Under this condition the incremental model is unstable in the linear sense and will respond with a positive exponential transient. Examination reveals, however, that the system will tend toward the correct position with a positive exponential and assume its proper stable relationship.

To study the manner in which the system responds in the time domain when the signal $V(s)$ is applied suddenly as a step function, the x and y outputs can be obtained by inversion of their respective transforms. Replacing $V(s)$ by V/s and letting $T=K_y+K_x|K_R|M$, equations 23 and 24 can be rewritten as

$$X_c(s)=\frac{K_x K_y}{T}\,\frac{V}{s^2}+\frac{M|K_R|K_x^2}{T}\,\frac{V}{s(s+T)} \qquad (26)$$

$$Y_c(s)=\frac{K_x K_y K_R}{T}\,\frac{V}{s^2}-\frac{K_x K_R K_y}{T}\,\frac{V}{s(s+T)} \qquad (27)$$

The reason for separating the transforms in this manner is seen by referring to Fig. 6 where the time functions $y_c(t)$ and $x_c(t)$ are sketched. It is seen that the two ramp functions will combine to yield perfect response but that the other two functions generate the transient and following errors. A plot of y_c versus x_c from these results is obtained by combining the time functions as shown in Fig. 7 where the two components of the steady-state error are shown. The following error e_{yss} is easily obtained from these components and is seen to agree with the result given in the preceding section.

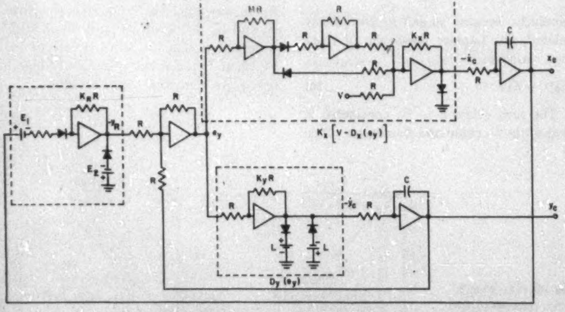


Fig. 8. Analog computer simulation of 2-dimensional control system

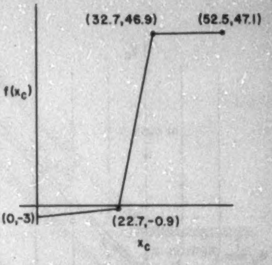


Fig. 9. $f(x_c)$ used in analog computer studies

Design of System

As in any control system the design problem is to choose certain system parameters such that the desired performance specifications are met. In the first place, it is assumed that there is a specification which limits the maximum speed in both the x- and y-directions. A maximum allowable error in the y-direction is also specified. The maximum speed in the y-direction sets the value of L, while the maximum speed in the x-direction gives $K_x V$. The maximum error specification is inserted by making $e_{y\max}=V/M$. Another consideration mentioned previously is that to obtain good dynamic performance it is desirable to utilize as large a gain K_y as possible in the y control loop. Increasing K_y will also reduce the steady-state following error until a point is reached at which

$$K_y=\frac{L}{\dfrac{V}{M}-\dfrac{L}{K_x K_R M}} \qquad (28)$$

Beyond this point further increases in K_y will reduce the following error no more because, due to the y-limiter, the speed in the y-direction cannot become greater than L regardless of the y-loop gain. It is advisable therefore to choose K_y

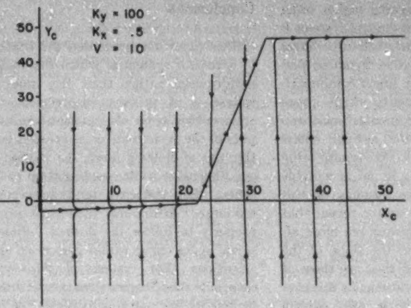


Fig. 10. Phase-plane trajectories obtained on analog computer

equal to or greater than the value given by equation 28, for the largest K_y encountered, once the other parameters have been set. This will insure the minimum possible following error consistent with the other limitations on the system. On the other hand, regardless of settings of K_y and other coefficients, the following error can never exceed the maximum value e_{ymax}.

Whether or not the dynamic performance of the system based on these considerations is satisfactory under all conditions, will of course be verified by analog computer studies.

Analog Computer Studies of System

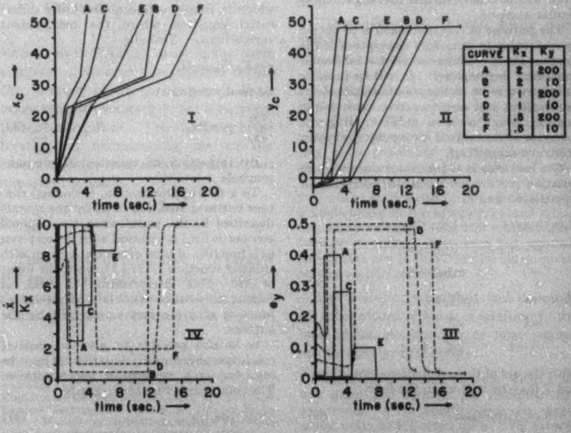
In order to verify the conclusions of the preceding sections, the system of Fig. 2 was programmed on an analog computer with $G_x(s)$ and $G_y(s)$ set at $1/s$. Performance curves were taken for various combinations of y-loop gain and maximum x velocity. The actual computer block diagram is shown in Fig. 8 and the functional relation $f(x_c)$ which results is shown in Fig. 9. The small slope for $0 \leq x_c \leq 22.7$ is due to the finite back resistance of the series diode in the $f(x_c)$ amplifier. There is also a negligibly small slope apparent for $32.7 \leq x_c \leq 52.5$ which is due to the small but nonzero forward resistance of the diode shunting the output of the $f(x_c)$ amplifier. As a practical consideration it is noted that wherever a diode is used in the series arm of an operational amplifier, the value of R for resistors associated with that amplifier should be small compared to the back resistance of the diode. This is done to minimize the effect of the finite back resistance of the diodes.

The value of e_{ymax} was chosen to be equal to 0.5, V was set to 10, and $\dot{y}_{max}$ was limited to about 22 volts per second. Fig. 10 shows the actual phase portrait of the system for $K_x = 0.5$ and $K_y = 100$. It was obtained by recording the x and y outputs of the computer on an x-y plotter for various values of initial conditions on the x and y integrators. The results graphically confirm the conclusions of the preceding section. The nonvertical slopes of the trajectories approaching the curve are due to the fact that $\dot{x}$ is not limited exactly at zero, as was explained previously, because of the forward resistance of the shunting diode on the x-channel of the command distributor. The actual curvature of the trajectories in the linear region is not very apparent since this region is an extremely narrow band about the curve. Over-all operation of the system, however, is exactly as predicted.

Time responses of x_c, y_c, e_y, and $\dot{x}/K_x$ were also obtained and are shown in Fig. 11. These curves show how the system slows down its x motion when the slope of $f(x_c)$ becomes large and speeds up again when the system enters a small slope region. Curve B, taken for a large $\dot{x}_{max}$ and small K_y, illustrates this effect most markedly. When the system enters the steep slope portion of $f(x_c)$ the error e_y becomes quite large, being almost equal to e_{ymax}, and the x velocity decreases considerably until the system leaves the region. At the other extreme, curve E, taken for small $\dot{x}_{max}$ and large K_y, shows only a slight change in x velocity and a small error in following the steep slope. These curves also show quite well the effect that the maximum x velocity $(\dot{x}_{max} = K_x V)$ has on the following error. It is noted that for the same value of K_y, reducing $\dot{x}_{max}$ decreases the following error, and that this effect is most marked for large K_y.

Design of Higher Order Systems

The preceding discussions relate to relatively simple systems in which the plant is represented as a simple integrator. With this assumption, the system is of second order and a phase plane representation combined with a linear analysis for small errors is possible. When a more complex plant such as a practical motor for which the transfer function $G(s)$ is $1/s(s+a)$, the equations become of fourth order and a phase-space representation must be used. To describe such a system would take four dimensions and it would be very difficult to visualize



CURVE	K_x	K_y
A	2	200
B	2	10
C	1	200
D	1	10
E	.5	200
F	.5	10

Fig. 11. Time-response curves obtained on analog computer

or sketch trajectories in this higher dimensional space. It is evident, therefore, that phase-plane (or space) sketches which served to guide the analysis of the simple system cannot be made for more complex systems.

Fortunately, it is possible to confirm that for any plant containing a single integration and additional dynamic effects consisting of a number of simple cascaded time constants, the system will always approach and enter the linear region surrounding the zero error condition for any initial conditions in the phase space. On the other hand, it is not possible fully to guarantee that there will not be some region which will cause the system to approach limit cycles or other forms of irregular behavior. To verify the behavior of the system under all circumstances requires exhaustive study on an analog computer very similar to that carried out for the simple system.

A recommended procedure, therefore, for studying and designing higher order 2-dimensional feedback control systems is to make a study in the small error linear region to synthesize good dynamic performance and to design linear compensating networks or elements which insure stability and good dynamical performance. Thus, when the system enters the linear region it will recover and bring the error toward zero in an acceptable manner. For instance, studies of this type with a practical motor reveal that the dynamics of the system are more affected by variations in the slope of the function to be followed than are those of the simple linear integrator. In this case the use of an automatic gain control as discussed which cancels out the effect of different slopes in the function $f(x_c)$ will probably be necessary. The linear analyses combined with studies on an analog or digital computer can thus serve as a basis for synthesis and design of 2-dimensional servomechanisms.

Conclusions

This paper has presented the concept of a control system in which the spatial performance rather than the time response is of primary importance. To achieve this type of operation the outputs of the system are cross-coupled with the aim of slowing down the motion in one dimension while speeding it up in the other, when the error tends to become too large. In this way both loops adjust properly to follow the desired 2-dimensional space curve within prescribed specifications. For systems of this type, especially those higher than second order, techniques are not now available for full analytic treatment but a combination of an approximate linear small signal approach with computer studies is required. Systems of this type show considerable promise in applications such as contour-cutting-tool controls which program themselves logically.

◆

Discussion

R. E. Kalman (Columbia University, New York, N. Y.): Although there is now available an extensive theory for the analysis of nonlinear control systems, theoretical considerations provide only the most meager guide for the synthesis of such systems at the present time. The problem of synthesis can be successfully approached only by a combination of physical insight and careful analysis, supported by a large amount of ingenuity and experience. This is just what the authors of this very fine paper have done. It is hoped that their pioneering work will stimulate further investigations of similar systems.

The purpose of this discussion is to obtain an approximate linearized description of the system in a manner somewhat different from that in the paper. It will be possible to improve some of the results obtained by the authors and also to arrive at a simple analogy which aids in understanding the behavior of the system when subjected to an arbitrary input $f(x_c)$.

The first step is to introduce a normalized notation for the nonlinearities occurring in equations 6 and 7. Define

$$
\begin{aligned}
N_1(x) &= x+1 && -1\leq x\leq 0\\
&= -x+1 && 0\leq x\leq 1\\
&= 0 && \text{otherwise}\\
N_2(x) &= x && |x|\leq 1\\
&= 1 && x\geq 1\\
&= -1 && x\leq -1
\end{aligned}
$$

With the aid of these definitions, equations 6 and 7 may be written in the form:

$$\dot{x}_c = K_x V N_1(e_y/e_{y\max}) \tag{6A}$$

$$\dot{y}_c = L N_2(e_y/e_{y\max}) \tag{7B}$$

where e_y is as defined by equations 1 and 5; K_x, V, L, $e_{y\max}$ are real positive constants as in the paper. Eliminating the time, we can write

$$\frac{dy_c}{dx_c}=\frac{\dot{y}_c}{\dot{x}_c}=\frac{L}{K_x V}\frac{N_2(e_y/e_{y\max})}{N_1(e_y/e_{y\max})} \tag{29}$$

This shows that dy/dx is a nonlinear function of $e_y/e_{y\max}$. If $|e_y/e_{y\max}|$ is small, we may approximate this function by multiplying its argument by its slope at the origin which leads to

$$\frac{dy_c}{dx_c}=\frac{L}{K_x V e_{y\max}}\,(f(x_c)-y_c)$$

which is simply a linear, first-order differential equation where the independent variable is x_c. Defining

$$\tau = K_x V e_{y\max}/L$$

we finally arrive at

$$\tau\frac{dy_c}{dx_c}+y_c=f(x_c) \tag{30}$$

By inspection of equation 30 we may conclude:

To a first approximation, the actual contour cut as a function of x_c by the system described in the paper when the required contour is $f(x_c)$ is identical with the response as a function of time of a linear system with transfer function $1/(\tau s+1)$, when the input is $f(t)$. This interpretation is valid, of course, only when $f(x_c)$ is a single-valued function of x_c—a fact emphasized by the authors.

It is also possible to give an explicit condition under which equation 30 may be regarded as a permissible approximation. The condition is that the quantity

$$|df(x_c)/dx_c|\,.\,\tau \tag{31}$$

must be sufficiently small. This follows by expanding the forced response of equation 30 in a series which depends on various derivatives of $f(x_c)$. See, for instance, reference 1 for additional details.

As a check, assume that $f(x_c) = K_R x_c$ in equation 30. Then $f(x_c)-y_c(x_c)\rightarrow y_{css}= K_R\tau$ as $x_c\rightarrow\infty$. Thus if $K_R\tau$ is sufficiently small, then the expression for the steady-state error y_{css} given by equation 13 of the paper may be seen to reduce to $K_R\tau$.

REFERENCE

1. CLOSED EXPANSION OF THE CONVOLUTION INTEGRAL, E. Arthurs, L. H. Martin. *Journal of Applied Physics*, New York, N. Y., vol. 26, 1955, p. 58.

Jack Sklansky (RCA Laboratories, Princeton, N. J.): This paper, in my opinion, is an important contribution to the control system art, because it strikes out in a new direction—or new "dimension," to venture a pun. The direction is new, because little similar material of this nature has previously been published, with the exception of some of the specialized techniques described in the unclassified literature on guided missiles.

The field of 2-dimensional control systems, one should note, covers more than the control of space variables; it applies to the control of any two time-dependent variables having some desired functional relationship, with time serving merely as a common dependent parameter. The following is a hypothetical application in which y is not a space variable:

x = distance of a semiconductor crystal above the surface of a melt
y = concentration of n- or p-type impurity in the melt (for producing p-n junctions in semiconductor crystals)

Thus, 2-dimensional control systems are a new class of systems in which time dependence takes on a minor role relative to intervariable dependence. A still further extension to the control-system concept can be made: namely, the class of control systems in which both time dependence and intervariable dependence are of primary importance. An example of this is the following:

x = density of vehicles on road X
y = density of vehicles on road Y (for controlling traffic on two parallel roads leading to a common destination, such as a bridge or tunnel)

Thus, three classes of control systems are now apparent: (1) systems in which the time dependence of one or more variables is of primary importance, and in which intervariable dependence is of secondary or no importance; (2) systems in which the dependence between two or more variables is important, time dependence unimportant; (3) systems in which time dependence and intervariable dependence are equally important. It is suggested that these three classes of systems be called "wave control systems," "path control systems," and "trace control systems," respectively, where wave, path, and trace denote in general multidimensional curves for which time dependence and intervariable dependence have appropriate amounts of importance.

A technical correction is suggested as follows:

From equation 8 it is clear that the authors intend to extend the concept of the phase plane to any two variables, rather than just distance and time. In the conventional phase plane the co-ordinates are position and velocity; hence, in the generalized phase plane the co-ordinates should be y and dy/dx. However, the authors plot curves of y versus x, and call them "phase-plane trajectories." It seems to me that they would be more properly denoted as "trajectories in the xy-plane."

Philip Sarachik and J. R. Ragazzini: We wish to thank Mr. Kalman and Dr. Sklansky for their interesting discussion of this paper.

Mr. Kalman's method of obtaining the system behavior for any specified single-valued functional relation $f(x_e)$ is quite novel. He shows that the behavior is equivalent to that of a system with a single "time constant" responding to an input $f(t)$. It is important to emphasize, however, that the approximation made in obtaining the simple result given in equation 30 is permissible only in the very restrictive condition given in equation 31. Actually the system will perform well within specifications even if this condition is not satisfied and so most systems of this type will not be designed to satisfy the condition. Unfortunately, therefore, in most systems we would still have to resort to an analog computer to see how the system performs with arbitrary functions $f(x_e)$. The single-time-constant analogy, however, is a good one and enables one to understand more clearly how the important system parameters affect the behavior of the system.

We cannot agree with Dr. Sklansky's objection to the use of the phrase "phase plane" in connection with the system discussed in our paper. There seems to be some confusion as to just what a phase plane, or, more generally, a phase space or state space, is. It is admitted that it has become common practice to use position and velocity (y, $\dot{y}$) as the variables in the phase plane. This does not mean, however, that these are the only permissible variables. The phase plane may be defined by any two variables which uniquely specify the state of the system and in some cases different variables may be preferable or even necessary. The specific justification for using "phase plane" in referring to the x-y plane of the system can be found by referring to reference 1. Here the authors in discussing systems described by two equations of the first order [$\dot{x} = P(x,y)$; $\dot{y} = Q(x,y)$] state, "In many cases the x-y plane is a phase plane and so we shall continue to call it that."

REFERENCE

1. THEORY OF OSCILLATIONS (book), A. A. Andronow, C. E. Chiaken. Princeton University Press, Princeton, N. J., 1949, p. 182.

Single-Phase Versus 3-Phase Service for Residential Air Conditioning

A. S. ANDERSON
FELLOW AIEE

CHASE HUTCHINSON
FELLOW AIEE

RESIDENTIAL air conditioning has made a significant contribution to the growth of loads in the warmer parts of the United States. Centrally located units for summer and year-round air conditioning appear on the horizon as potential loads of unprecedented magnitude to nearly all parts of the country. All phases of utility operations will be affected but of immediate concern is the determination of the characteristics of the distribution system. The first question to be answered is: "Shall single-phase or 3-phase service be provided?" The customer, as well as the company, is vitally interested in the answer.

The most prevalent sizes of motors will be 2, 3, and 5 hp (horsepower). In the past, motors of these sizes were seldom used in residences and the choice between single phase and 3-phase was not of great economic importance. Therefore, the practice varies between companies and sometimes in different areas of the same company. Now however, utilities realize that a study of the problem on a long-range basis is necessary to set a basis for present-day policy. The company, customer, and manufacturer are equally concerned with over-all economics of this important problem.

Purpose

The purpose of this paper is to present the results of a study made in 1954 to determine the difference in cost of a single-phase system as compared to a 3-phase system for serving 3-hp and 5-hp air-conditioning units and heat pumps. Customer costs as well as utility costs are considered.

The secondary systems that are compared are as follows:

1. Single-phase 120/240-volt 3-wire system.
2. Three-phase 120/240-volt 4-wire delta system.

Note that the 3-phase 120/208-volt system is not economical for general overhead residential secondaries and, while considered in the study, data concerning it will not be presented. Higher secondary voltages are now being considered by the industry generally as a possibility in the future with greater load densities for residential areas. The cost of single-phase service as compared to 3-phase service will be more favorable to a single-phase service at any proposed higher secondary voltages. For this reason higher voltage secondary costs were not studied, since costs are in favor of the single-phase service at 120/240 volts.

The following situations and demands for other electrical loads in addition to the load of air conditioners or heat pumps are covered by this study:

1. Existing secondary system with 1.25-kva diversified demand per customer and saturation of air-conditioning or heat-pump units in 4% steps up to 50% saturation.
2. New secondary system in new residen-

Paper 57-6, recommended by the AIEE Domestic and Commercial Applications Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted August 23, 1956; made available for printing October 22, 1956.

A. S. ANDERSON and CHASE HUTCHINSON are with Ebasco Services, Inc., New York, N. Y.

tial development with 1.25-kva diversified demand per customer and saturation of air-conditioning or heat-pump units of 25, 50, and 100%.

3. New secondary system in new residential development with 7.5-kva maximum demand per customer and saturation of air-conditioning or heat-pump units of 100%.

The following individual customer demands for the air-conditioning or heat-pump loads were assumed.

Unit	Single Phase, Kva	3 Phase, Kva
3-hp air conditioning	4.65	5.05
5-hp air conditioning	8.31	8.74
3-hp heat pump	9.88	10.10
5-hp heat pump	16.80	17.10

The 3-hp heat pump includes 6 kw of auxiliary resistance heating units and the 5-hp heat pump includes 10 kw.

Conclusions

1. The sum of utility investment and customer investment is less with the single-phase 120/240-volt system than with the 3-phase 240-volt system for 3- and 5-hp motor-driven compressor units. From data developed, indications are that savings would also be realized for single-phase motor units of larger capacity.

2. The utility investment in a single-phase system including primary, transformers, secondary, meters, and services,

Table I. Average Utility Investment Savings Per Unit for Single-Phase 120/240-Volt System as Compared to 3-Phase 120/240-Volt Delta System

	Existing Residential System 1.25-Kva Demand for Other Load, Saturation		Newly Constructed Residential System 1.25-Kva Demand for Other Load, Saturation		
Unit	4%	50%	25%	50%	100%
3-hp air conditioner	$343	$ 92	$174	$105	$69
5-hp air conditioner	377	133	167	111	73
3-hp heat pump	307	81	123	71	65
5-hp heat pump	396	121	109	79	74

Table II. Approximate Utility Investment Savings Per Unit for Single-Phase 120/240-Volt System as Compared to 3-Phase 120/240-Volt Delta System With 100% Saturation of Air-Conditioner Units in a New Residential Development

Unit	Transformers	Meters	Other	Tota
3-hp air conditioner	$42	$35	$5	$ 82
5-hp air conditioner	77	35	−1	111

Table III. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 3-Hp Air-Conditioning Units Added to Existing 7,200 to 120/240-Volt Single-Phase System*

Saturation		Savings Per Unit Averaged For Saturation Shown†				
No.	%	Transformers	Primary	Secondary	Meters	Total‡
6	4	$183	$117	$10	$35	$343
12	8	157	57	7	35	257
18	12	100	38	6	35	180
24	17	121	25	5	35	187
30	21	88	20	5	35	149
36	25	81	14	4	35	135
42	29	77	12	4	35	129
48	33	59	11	4	35	110
54	37	50	8	4	35	98
60	42	48	7	4	35	95
66	46	46	6	4	35	92
72	50	48	5	3	35	92

* Existing system consists of 25-kva transformer with six spans per transformer. Secondary wire size, American Wire Gauge no. (AWG) 4 copper. Load per customer, 1.25 kva. four customers per pole. Air-conditioning unit installed not more than one span from transformer to limit flicker. Minimum size transformer for "lighting" load is 15 kva. Maximum load per transformer, 120%. Open Y to open-delta banks used for 3-phase service. Load for single-phase unit including fans equals 4.6 kva. Load for 3-phase units including fans equals 5.0 kva.

† For each saturation shown the average costs are those which will most likely have resulted by doing those things required for each lesser saturation progressively up to the saturation shown.

‡ Includes excess cost of 3-phase service drop over single-phase service drop of $1.

Table IV. Investment Savings of Single-Phase Over 3 Phase per Installation for Utility System to Serve 5-Hp Air-Conditioning Units Added to Existing 7,200 to 120/240-Volt Single-Phase System*

Saturation		Savings Per Unit Averaged For Saturation Shown†				
No.	%	Transformers	Primary	Secondary	Meters	Total‡
6	4	$222	$112	$7	$35	$377
12	8	249	59	7	35	351
18	12	153	37	6	35	232
24	17	123	25	5	35	189
30	21	112	17	5	35	170
36	25	115	13	4	35	168
42	29	112	11	4	35	163
48	33	100	9	4	35	149
54	37	96	7	4	35	143
60	42	97	6	4	35	143
66	46	92	5	4	35	137
72	50	89	5	3	35	133

* Existing system consists of 25-kva transformer with six spans per transformer. Secondary wire size, AWG *4* copper. Load per customer, 1.25 kva. Four customers per pole. Air-conditioning unit installed not more than one span from transformer to limit flicker. Minimum size transformer is 15 kva for "lighting" load for 3-phase secondary and 25 kva for single-phase secondary. Maximum load per transformer, 120%. Open-Y to open-delta banks used for 3-phase service. Load for single-phase unit including fans equals 8.3 kva. Load for 3-phase units including fans equals 8.7 kva.

†‡ See corresponding footnotes, Table III.

Table V. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 3-Hp Heat Pumps Added to Existing 7,200 to 120/240-Volt System*

Saturation		Savings Per Unit Averaged For Saturation Shown†				
No.	%	Transformers	Primary	Secondary	Meters	Total‡
6	4	$144	$117	$10	$35	$307
12	8	129	57	10	35	232
18	12	80	26	9	35	151
24	17	83	22	8	35	149
30	21	86	15	7	35	144
36	25	65	12	6	35	119
42	29	68	10	6	35	120
48	33	47	7	6	35	96
54	37	38	7	6	35	87
60	42	32	5	6	35	79
66	46	36	3	5	35	80
72	50	37	3	5	35	81

* Existing system consists of 25-kva transformer with six spans per transformer. Secondary wire size, AWG *2* copper. Load per customer, 1.25 kva. Four customers per pole. Heat pump installed not more than one span from transformer to limit flicker. Minimum size transformer is 15 kva for "lighting" load. Maximum load per transformer, 150%. Open-Y to open-de'ta banks used for 3-phase service. Load per unit, including fans and 6-kw supplemental heating, equals 9.9 kva for single phase and 10.1 kva for 3 phase.

†‡ See corresponding footnotes, Table III.

is lower than the investment in a 3-phase system as shown in Tables III to X and further summarized in Table I.

Tables XI and XII show the possible savings in utility investment for a new residential area where the saturation of units is 100% and the demand of the load other than air conditioning is 7.5 kva. It may be seen that most of these savings are in transformer costs and meters and may be summarized as shown in Table II.

3. The extra cost to the customer of single-phase units over 3-phase units of 3-hp and 5-hp ratings should eventually approach about $50 and $80 respectively. At present the differentials are somewhat greater for units of some manufacturers. This greater differential cost is believed due to differences in assembly techniques, pricing, and methods, and relative production.

The extra cost of single phase is caused by need for starting and running capacitors and starting relays. If the running power factor of 3-phase motors were made equal to running power factor of single-phase motors, the cost of the 3-phase unit would increase.

4. Estimates of the cost of residential wiring for single-phase and 3-phase air-conditioning units of 3-hp and 5-hp ratings were received from ten operating utility

Table VI. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 5-Hp Heat Pumps Added to Existing 7,200 to 120/240-Volt System*

Saturation		Savings Per Unit Averaged For Saturation Shown†				
No.	%	Transformers	Primary	Secondary	Meters	Total‡
6	4	$228	$111	$10	$35	$396
12	8	168	44	10	35	269
18	12	120	19	9	35	195
24	17	105	15	8	35	175
30	21	88	10	7	35	152
36	25	78	6	6	35	137
42	29	64	4	6	35	121
48	33	75	3	6	35	131
54	37	76	2	6	35	131
60	42	72	1	6	35	126
66	46	64	1	5	35	117
72	50	68	1	5	35	121

* Existing system consists of 25-kva transformer with six spans per transformer. Secondary wire size, AWG *2* copper. Load per customer, 1.25 kva. Four customers per pole. Heat pump installed not more than one span from transformer to limit flicker. Minimum size transformer for "lighting" load is 25 kva. Maximum load per transformer, 150%. Open-Y to open-delta banks used for 3-phase service. Load per unit, including fans and 10-kw supplemental heating, equals 16.8 kva for single phase and 17.1 kva for 3 phase.

† See corresponding footnote, Table III.

‡ Includes excess cost of 3-phase service drop over single-phase service drop of $12.

Table VII. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 3-Hp Air-Conditioning Units in New Areas*

Saturation, %	Transformer	Primary	Secondary	Meters	Total†
25	$115	$18	$5	$35	$174
50	61	6	2	35	105
100	30	2	1	35	69

* Three-phase system is served by open-Y to open-delta transformer banks. Light and appliance load assumed to be 1.25 kva per customer at transformer. Air-conditioning load, including fans, equals 4.6 kva for single phase and 5.0 kva for 3 phase.

† Includes $1 for extra cost of 3-phase over single-phase service.

Table VIII. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 5-Hp Air-Conditioning Units in New Areas*

Saturation, %	Transformer	Primary	Secondary	Meters	Total†
25	$113	$13	$5	$35	$167
50	69	4	2	35	111
100	35	1	1	35	73

* Three-phase system is served by open-Y to open-delta transformer banks. Light and appliance load assumed to be 1.25 kva per customer at transformer. Air-conditioning load, including fans, equals 8.3 kva for single phase and 8.7 kva for 3 phase.

† See corresponding footnote, Table VII.

Table IX. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 3-Hp Heat Pumps in New Areas*

Saturation, %	Transformer	Primary	Secondary	Meters	Total†
25	$68	$13	$6	$35	$123
50	28	4	3	35	71
100	26	1	2	35	65

* Three-phase system is served by open-Y to open-delta transformer banks. Light and appliance load assumed to be 1.25 kva at the transformer. Heat-pump load, including fans and 6.0-kw supplemental heating, equals 9.9 kva for single phase and 10.1 kva for 3 phase.

† See corresponding footnote, Table VII.

Table X. Investment Savings of Single Phase Over 3 Phase per Installation for Utility System to Serve 5-Hp Heat Pumps in New Areas*

Saturation, %	Transformer	Primary	Secondary	Meters	Total†
25	$63	$9	$7	$35	$109
50	44	1	4	35	79
100	41	1	2	35	74

* Three-phase system is served by open-Y to open-delta transformer banks. Light and appliance load assumed to be 1.25 kva at the transformer. Heat-pump load, including fans and 10-kw supplemental heating, equals 16.8 kva for single phase and 17.1 kva for 3 phase.

† Includes $5 for extra cost of single-phase over 3-phase service.

Table XI. Savings in Investment and Capitalized Loss of Single Phase Over 3 Phase per Installation in Utility System to Serve a 100% Saturation of 3-Hp Air-Conditioning Unit in a New Area With 7.5-Kva Demand per Customer for Other Load*

	Customers per Transformer		
	8	12	16
Transformers	$24	$42	$44
Secondary	1	2	
Service	1	1	1
Meters	35	35	35
Subtotal	$61	$80	$80
Capitalized Losses	13	2	7
Total	$74	$82	$87

* Three-phase system is 4-wire delta.

Table XII. Savings in Investment and Capitalized Loss of Single Phase Over 3 Phase per Installation in Utility System to Serve a 100% Saturation of 5-Hp Heat Pumps in a New Area With 7.5-Kva Demand per Customer for Other Load*

	Customers per Transformer		
	4	8	12
Transformer	$ 91	$ 60	$ 77
Secondary		0	0
Service	(5)	(5)	(5)
Meters	35	35	35
Subtotal	$121	$ 90	$107
Capitalized Losses	15	15	4
Total	$136	$105	$111

( ) Denotes less cost for 3 phase.
* Three-phase system is 4-wire delta.

companies. Generally, the difference in cost between wiring for single-phase units or 3-phase units does not depend upon whether additional wiring for an existing residence or entirely new wiring for a new residence is being considered. Sixteen sets of estimates indicate that wiring costs for single-phase units are lower by a range of from $6 to $77 or an average of $48. Five sets of estimates show that wiring for 3-phase units are lower by a range of from $3 to $38 or an average of $16. The difference in cost was not affected appreciably by whether the motor is 3-hp or 5-hp rating. Three sets of estimates for 5-hp heat pumps with 10-kw auxiliary heating were received. Two sets show the wiring cost for single phase units to cost less by an average of $23. The other set shows 3-phase wiring to cost less by $7. The difference in wiring cost for single-phase and 3-phase units generally favors the use of single-phase units. The average of all differences is $30 in favor of single phase.

5. Capacitor-start, capacitor-run single-phase motors compare very favorably in operating characteristics with 3-phase induction motors in 3-hp and 5-hp sizes. Some of comparative characteristics of one make of 5-hp motor are:

	Single Phase	3 Phase
Rated voltage	230	220
Running power factor, %	93	83
Starting power factor, %	87	69
Locked rotor amperes	100	90
Efficiency, %	80	84
Locked rotor torque, foot-pounds	36	28

◆

Discussion

Robert F. Lawrence (Westinghouse Electric Corporation, East Pittsburgh, Pa.): We are in general agreement with the authors' results of the very extensive study which they conducted. I would like to discuss the basis of our agreement and also would like to offer an alternate method of presentation of data. There is, too, one exception to their general conclusions; namely, that while 3 phase is more costly than single phase at today's conventional secondary voltages, it is not necessarily true at a higher secondary voltage.

In our studies of higher secondary voltage distribution systems for residential areas,[1,2] the cost of a 3-phase 120/208-volt 4-wire system was determined and compared to the cost of a single-phase 120/240-volt system. The primary phase-to-phase voltage for both systems was 34.5 kv. The primary of the 3-phase system utilized three conductors.

Because our studies applied to new areas our comments are confined to the results contained in Tables VII, VIII, IX, and X.

For purposes of orientation, let us compare the 34.5-kv single-phase and 3-phase systems that we studied. Later I will make an allowance for costs of a 12.5-kv primary voltage.

In the 120/208-volt 3-phase Y secondary system and the open-Y 3-phase 120/240-volt 4-wire delta system, the number of primary conductors is the same. The size of primary conductors for a given primary voltage would also be the same. The distribution transformers would be 3-phase or three single-phase units in the Y system and two units, open-Y, open delta, in the delta system. All of these factors affect costs and I would expect the costs to be slightly lower had we studied the 120/240-volt 3-phase 4-wire delta system.

The total dollar investment savings, exclusive of meters, of the single-phase system over the 3-phase system is practically constant at about $110 per house over a range of load densities of 1.6, 2.5, and 5.0 kva per house. The kilovolt-amperage per house is the average diversified peak demand per house as viewed for 1,000 houses. The total savings at 10 kva per house is approximately $200. This savings might be expected to increase at the higher density although we believe that this instance is attributable to the costs used for the larger 19.9-kv transformers to supply the higher density load.

All of our studies were based on a simultaneous economic analysis of the primary, distribution transformer, and the secondary, to obtain optimized system costs. It is a lengthy and tedious process and would be very time-consuming to repeat for a primary voltage other than 34.5 kv. However, we have been able to make some quick but reliable approximations of costs of a primary voltage of 12.47 kv to compare our results more directly with the authors' studies. From these, we find that the total investment savings per customer of a single-phase system over a 3-phase system is approximately $55 over a range of load densities of 1.5, 2.5, 5.0, and 10.0 kva. The savings include investment on primaries, distribution transformers, and secondaries, but not meters. The savings of $35 per meter as used by the authors could be added to the foregoing figure to give a total investment savings of $90 per house.

In Tables VII–X, the investment savings are given per installation, which we presume to mean per air-conditioning or per heat-pump installation. Thus, in the case of 25% saturation, only one house in four has an air conditioner and to determine the savings per customer the values given for the savings of the transformer, primary, and secondary should be divided by 4. Similarly for 50% saturation, these costs should be divided by 2. We have for Table VII:

% Saturation	Savings per Customer
25	$69.55
50	69.50
100	69.00

Repeating the foregoing for Tables VIII, IX, and X gives similar results, namely, that the dollar savings is practically constant. The value of $90 arrived at is in remarkably close agreement considering the inevitable variations in studies by different groups. We believe that this is the way these data should be presented.

It is our thought, however, that if the investment savings is to be given on some basis other than per customer, the following evaluation is more correct than the per-installation basis shown.

We have selected Table VII to illustrate the point. It is assumed that 12 customers are served per distribution transformer. The savings will be apportioned on the basis of the customer kilovolt-amperage demand.

Let X = dollar savings per kilovolt-amperage of demand

1.25-kva demand for non-air-conditioning customer

6.25-kva demand for air-conditioning customer

then, for 25% saturation:

$$(1.25\times 9)X+(6.25\times 3)X=\left(\frac{115+18+5}{4}\right)\times 12$$

$= \$13.8$ savings per kilovolt-amperage

$\$13.8\times 1.25=\$17.25$ savings per non-air-conditioning customer

$13.8\times 6.25=\$86.25$ savings per air-conditioning customer

$86.25

+35.00 meter savings

$121.25 total savings per air-conditioning customer

−17.25 savings of 1.25-kva load that is not air conditioning

$104.00 incremental savings per air-conditioning customer

For 50% saturation, we have

$ 9.20 savings per kilovolt-amperage

11.4 savings per non-air-conditioning customer

$57.0 savings per air-conditioning customer

+35.0 meter savings

$92.0 total savings per air-conditioning customer

−11.4 savings of 0.125-kva load that is not air conditioning

$80.6 incremental savings per air-conditioning customer

The foregoing method appears to give exhorbitant values of savings and costs to serve the high-demand customer. Since the total costs of the single-phase and 3-phase systems are not given in the paper, we cannot properly evaluate their costs.

I mentioned earlier the costs of the single-phase and 3-phase secondary systems at the higher secondary voltage. Our studies showed them to be about a standoff, exclusive of metering. Three-phase metering is more costly, but with single-phase service to each home, this would no longer be a problem. We still conclude that single- or 3-phase higher secondary voltage distribution are about equal in cost, and we are inclined to favor the 3-phase system.

References

1. Residential Distribution—An Analysis of Systems to Serve Expanding Loads, S. B. Griscom, R. F. Lawrence. *AIEE Transactions*, vol. 75, pt. III, Aug. 1956, pp. 533–42.

2. Economic Comparison of Secondary Voltages; Single- and Three-Phase Distribution for Residential Areas, R. A. Zimmerman, H. E. Lokay. *Ibid.*, pp. 543–52.

A. E. Diehl (Borg-Warner Corporation, York, Pa.): We would like to commend the authors of this paper on doing a very comprehensive job with respect to single-phase versus 3-phase service for residential air conditioning. We believe that this is an excellent, condensed summary of a complete study conducted by the authors.

This paper concerns the application of 3- and 5-ton air conditioners, both of the conventional and the heat-pump type. The authors are to be complimented on selecting the most popular sizes of residential air conditioners as a basis for their study.

The results, as summarized, are quite indicative that definite savings to the utility result by furnishing single-phase power for the various types of distribution systems that are generally being used today.

The authors are aware of the fact that the single-phase air conditioners in 3- and 5-ton sizes are considerably more expensive in that they involve the use of starting relays and starting and running capacitors. The additional electrical components, naturally, result in more field service than is normally experienced on polyphase applications. It is therefore necessary to compare the increased unit cost plus the increased possibility of field service cost versus increased power factor and possible savings in the electric wiring costs for the customer, before making general conclusions that would apply to all areas of the United States.

We believe that the air-conditioning and refrigeration industry requires further time and experience to develop less costly and more reliable single-phase equipment before the use of single phase only should be considered for application to residential air conditioning. Experience to date indicated that the polyphase units are more rugged and will withstand more abuse.

There is an added incentive to consider single-phase power supply more seriously by the air-conditioning and refrigeration manufacturers in view of a rather increasing difficulty experienced with single-phasing of polyphase circuits by electrical disturbances, power failures, etc. It is difficult to provide necessary motor protection for this condition, which, of course, is eliminated when single-phase supply is used.

May we take this opportunity to congratulate the authors on a fine, thorough study of a very pertinent problem with respect to the application of residential air conditioning and its inherent electrical application problems.

H. L. McClure (Delaware Power and Light Company, Wilmington, Dela.): Study of this paper raises a question of prime importance to those companies with 2,400/4,160-volt distribution. The comparisons made in the paper are for 7,200/12,470-volt primaries which have relatively high single-phase to neutral load capabilities. Would the same conclusions hold for 2,400/4,160-volt primaries, where presumably much more of the primary system would have to be multiphase?

We are greatly in favor of single-phase supply because of the cleaner appearance of lines and poles, simplified tree trimming, ease of banking secondaries, better power factor of motors, and, of course, lower cost.

Carl C. Ambrosius and **Robert H. Sarikas** (Illinois Power Company, Decatur, Ill.): The results obtained by the authors check fairly close with those we obtained in a similar study prepared in 1955. That study indicated a savings to the utility as high as $331 per 5-hp unit in 12.47-kv suburban areas where usually a single customer is served by a transformer. The savings with eight customers served from the same transformer on the 4.16-kv system were estimated to be $162 per customer. In general, it was found that the minimum increase in utility investment to provide 3-phase service was in the order of 30% of the required single-phase investment. The author's findings have also been substantiated by a number of other investigations.[1,2,3]

Customer equipment cost differences obtained from suppliers varied from $40 to as much as $125. The range of cost estimates for rewiring, received from contractors, was considerable. Three-phase wiring averaged about $25 higher than single-phase wiring. In many instances the total cost to the customer, equipment plus wiring, may be less for single-phase than for 3-phase service, especially if the existing single-phase service is adequate.

One argument frequently used by customers and contractors in favor of 3-phase service is increased reliability and efficiency. The tabulation in paragraph 5 of the author's "Conclusions" shows that the slight difference in efficiency is almost negligible. This tabulation also shows that with identical motor terminal voltages the starting current of a 3-phase motor is approximately equal to that of a single-phase motor. As a consequence, the flicker due to 3-phase motor starting is about equal to that of single-phase motor, and may be greater where the saturation of 3-phase motors permits the use of a smaller distribution transformer for the "lighting" load in an open-Y to open-delta transformer bank.

As regards reliability, a check with a large supplier of heat pumps and central air-conditioning units reveals that the failure rate of 3-phase units is about equal to that of single-phase unit. This is attributed primarily to motor burn-out caused by single-phasing due to an outage of one phase of the primary distribution system from such causes as fuse-blowing or single-phase oil-circuit recloser operation. A case of this type of burn-out, due to failure of the motor phase protection to operate, is described in a paper concerning some of the early heat-pump test installations.[4] Further, it would appear that in a residential area, the probability of an outage of either phase of an open-Y system would be considerably greater than that of an individual phase, due to the increased exposure in circuit-miles. Since the loss of either phase will result in a 3-phase service interruption, this matter should be of importance to both utility and customer, especially with regard to heat-pump applications.

References

1. Problems Posed by Home-Motor Loads, H. E. Campbell. *Electrical World*, New York, N. Y., May 17, 1954.

2. 120/240 V Can Serve Bigger Loads, H. G. Barnett, L. W. Manning. *Ibid.*

3. Serving Air-Conditioning Cheaply, H. W. Grate. *Ibid.*, Dec. 6, 1954.

4. Residential Heat Pump Experiments in Philadelphia—Installation and Operating Experience, J. H. Harlow, G. E. Klapper. *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 366–75.

Wayne H. Johnson (American Gas and Electric Service Corporation, New York, N. Y.): The conclusions reached in the paper that 3- and 5-hp motors and heat pumps can be more economically served single phase than 3 phase confirms our conclusions made over 10 years ago. A review of present costs indicates there is no reason to change our thinking from an economical cost to serve standpoint.

However, I do wish to comment specifically on the flicker limitation proposed in this paper. At present, the flicker problem in connection with heat pumps does not appear to be serious. Flicker difficulties in connection with automatically controlled heat pumps starting across the line are more likely to occur during equilibrium weather, that is, when the outside weather is close to that required by the inside controlling thermostats. Starting under these conditions may be as frequent as once in 15 to 20 minutes.

A review of the various flicker curves indicates that an average requirement permits about 4% flicker for a frequency of 15 to 20 minutes. The flicker to be expected at the secondary terminals of a 7,200/120/240-volt distribution transformer with sustained high side voltage when a 5-ton General Electric Weathertron is starting is 4% on a 25-kva, 2.4% on a $37^1/_2$-kva, and 1.67% on a 50-kva load. Adding a 125-foot span of no. *2* copper, or no. *1/0* steel-reinforced aluminum cable secondary on racks, causes these flicker percentages to become 8.13%, 6.59%, and 5.94% respectively. This indicates that other customers connected to the same service pole (one span from the transformer) might complain during equilibrium weather conditions.

We wish to point out, however, that even though our calculations indicate a possible difficulty, our experience does not prove that we have it. The matter of flicker, in connection with single-phase heat pumps, needs to be more completely explored and our flicker curves need to be re-examined.

William A. McGrath (Carrier Corporation, Syracuse, N. Y.): I am commenting on this paper as an air-conditioning manufacturer. From the standpoint of our industry, I think it is proper to say that the analysis offered represents a real contribution to our understanding of the electrical utility problems as influenced by the air-conditioning load. It is our feeling, however, that this is a subject which is not static and that conclusions that might be proper today could very well be obsolete another day.

There are a number of questions and observations which arise in the mind of the equipment manufacturer which may be worth offering for discussion.

1. While the title of the paper restricted the scope to residential air conditioning, it is hard to see how one can exclude commercial air conditioning from such a discussion since the equipment that is used for residential and commercial air conditioning is in the majority of cases identical. It would seem, therefore, that a complete answer to the question is not found until the needs of the commercial market are considered. If single-phase equipment is really the better choice for residential service, is this not also true for commercial service? Obviously, if the manufacturer's efforts could be directed to a standardized line of equipment, more progress in cost reduction and improvement would be achieved than if both types have to be furnished.

2. This paper states that the extra cost of single-phase air conditioners over 3 phase should be $50 for the 3-hp and $80 for the 5-hp size. It is unfortunate that the authors did not disclose the source of these estimates because the figures do not agree with our knowledge of the present cost level of this equipment. It is apparent from a review of current selling prices of different manufacturers in this industry that the actual cost will average $60 to $65 more for a 3-hp unit and $110 to $115 more for a 5-hp unit.

Also, in arriving at projected or ultimate cost differentials, it would be well to estimate the inherent excesses of materials and components that occur in the single-phase versus 3-phase equipment. Extra cost is represented in the capacitors required, starting relays, and equipment, and in addition more copper and iron are required in single-phase equipment for the same performance. Even beyond this, current single-phase motors are commonly machine-wound, which tends to reduce the labor cost, but most 3-phase motors of the same size are not machine-wound. On this basis, it could be assumed that if the usage of the 3-phase motors were increased the cost differential might favor the 3-phase equipment by an even greater margin.

3. A saving in utility investment cost was tabulated and indicates that 3-phase transformers and meters are substantially more expensive than single-phase units. It would be interesting to know why, for the same kilowatt capacity, this cost discrepancy should exist.

4. The discussion of the cost of residential wiring as between single-phase and 3-phase equipment does not actually present a very good case for assuming that these costs will favor single phase. The wide discrepancy of estimates received would seem to indicate that some factors other than the basic factor of cost and labor enter into these estimates. It would seem that a more precise comparison could be obtained by directly estimating and tabulating material costs and labor costs independently for the two systems. Also in this connection, it would be interesting to know what the estimated difference in utility wiring cost is between the two systems.

5. A related topic that might be appropriate for discussion in this connection is the matter of voltage regulations and phase unbalance on air-conditioning loads. It seems rather certain that air-conditioning equipment is frequently damaged by being exposed to voltages outside of the range of standards recommended in the Edison Electric Institute–National Electrical Manufacturers Association Standard *R-6*. Even beyond this, we have encountered serious conditions of phase unbalance which are not limited by this standard. In some cases these destructive conditions have been traced to the use of open-transformer arrangements, apparently inspired by the desire to place 3-phase equipment on single phase distribution systems. If a case can be made for better voltage regulation with the single-phase systems, this would be an important consideration to the industry.

6. Another similar question which might be worth exploring is the question of light flicker or television sag as might result from starting impulses. The question here is whether there is any margin of superiority of the 3-phase service versus single phase on this score.

B. M. Gallaher (Texas Electric Service Company, Fort Worth, Tex.): The paper is a well-documented contribution to the problem facing many utilities as to the choice of furnishing single-phase or 3-phase power for residential use. The utility with which I am connected supplies both single-phase and 3-phase service to many residential customers. It is the purpose of this discussion to emphasize and to concur with the authors' conclusions and possibly to view the problem from a slightly different angle. Most people in the electrical business take 3-phase primary distribution for granted but are inclined to overlook the costs of components necessary to obtain 3 phase at secondary voltage levels. Some of these components are:

1. Although the kilowatt-hours (or energy) in supplying power to home appliances is practically the same for either single-phase or 3-phase power, the principal part of the "cost to serve" is represented in cost of the demand component (fixed costs of plant investment) rather than in the energy (fuel) component, hence any major saving in investment cost is an economical solution.

2. One way to consider the cost of a distribution transformer, including its accessories of lightning arresters and cutouts, may be expressed as:

Total installed cost = base cost + $K \times$ kva

where "base cost" is that component of cost that is approximately constant for several sizes of transformers in the range used for distribution loads, K is the incremental cost per kilovolt-ampere of transformer capacity, and "kva" is the kilovolt-amperage nameplate rating of the transformer.

Typical installed costs for a 7,200-volt/120/240 transformer is in the order of:

Total installed cost = $210 + $10 × kva

When two transformers are installed to serve a 3-phase secondary, the costs become:

Total installed cost = $410 + $10 × kva

Obviously, this additional $200 in the "base cost" is the minimum additional cost to establish a 3-phase secondary source to serve a single customer. Should two or more 3-phase residences share the same secondary, then this extra cost may be prorated.

3. As brought out in the paper, the investment cost of 3-phase meters is considerably greater than the investment cost of single-phase meters. In addition to investment costs, higher operating costs, both in maintenance and testing of 3-phase meters, are appreciable.

4. Additional operating costs to serve a 3-phase customer must include provision for monitoring load on two transformers rather than a single transformer. With our present practice of relating kilowatt-hours to kilovolt-amperage size of transformer for monitoring loads on transformers this mixture of 3-phase and single-phase loads adds another factor. Voltage balance on 3-phase motors is quite critical, necessitating more attention being given to balancing distribution loads between each phase on the primary. Also, with capacitors used on the newer single-phase motors, secondary power factor has been increased, thereby lowering the reactive power requirements.

The Southwest has had a phenomenal increase in saturation of home air conditioners during the past 10 years. From a negligible saturation at the end of the war, it is estimated that the present saturation is in the order of 25%. The problem of single-phase service versus 3-phase service to homes is, therefore, a real problem, and a studied approach to it, such as the paper under discussion, is greatly appreciated.

H. E. Campbell (General Electric Company, Schenectady, N. Y.): This paper gives a very fair comparison of the over-all economics of serving single-phase versus 3-phase residential air conditioning. The conclusions reached in this analysis are very similar to other studies of this subject which we have seen.

The authors point out that the cost of single-phase service as compared to 3-phase service will be more favorable to a single-phase service at any proposed higher secondary voltages. Another Winter Meeting paper[1] analyzes the cost of the higher voltage secondary and it shows that the higher voltage single-phase is considerably less expensive than the higher voltage 3-phase system, even for customer loads ten times that of a present day "full-use customer." This certainly confirms the authors' comment.

Despite the fact that economic analyses show single phase to be the most economical from an over-all point of view, there are some utilities who still wish to supply these loads 3 phase. This is usually done so that the utility's customer can get a lower cost unit and thereby encourage him to add the electrical load. When we first looked at this problem the 3-phase 5-ton unit at the customer's level was about $200 less expensive than the single phase. At the time the authors analyzed the problem hermetically sealed motors had been applied to the units and the price difference was about $80 less for the 3-phase unit. I have just learned that there is a differential of only about $35 at the customer's level for our 1957 heat pumps. This just about balances the $30 wiring cost in favor of single phase that the authors found in their analysis. Under these conditions the customer should not care which service he receives.

When single-phase service is used, the utility gets a load with about 10% better power factor, no difficulty from voltage unbalance, and no single-phasing problem. The economics are also in favor of single phase now and indications are that it will be more so in the future. Seldom do we get an answer to an economic analysis as clearcut as this one.

REFERENCE

1. ECONOMIC ANALYSIS OF RESIDENTIAL SECONDARY SYSTEMS, H. E. Campbell, H. E. Sinnott. *AIEE Transactions*, vol. 76, pt. III, 1957 (*Paper no. 57-169*).

Bryce Brady and **Bela Mohr** (Oklahoma Gas and Electric Company, Oklahoma City, Okla.): The authors have selected a very timely subject. The years 1956 and 1957 will probably show a new record in sales of central air-conditioning units for residential use. The indications are that the trend toward air-to-air central air-conditioning units in residential areas will increase. Certainly the people of the utility industry are studying their distribution design so that this new load will be served in the most economical manner.

We agree with the results as given by the authors. Studies made by our company during the last 3 years also show that single-phase utilization in residential areas is more economical than 3-phase utilization for all conceivable load densities. Our studies show that the 3-phase units are a little less expensive for the customer but that the additional investment required by the utility is more than enough to offset any possible saving to the customer. The 3-phase distribution system utilizing the open-delta transformer bank design as used by the authors is the most economical for 3-phase distribution. It gives every possible economic benefit to 3-phase utilization.

The economic analysis does not give credit to the single-phase system for the higher power factor of single-phase capacitor-type motors. The 3-phase system should be penalized by the addition of sufficient capacitors to produce equal power factor.

The extra cost to the customer for single-phase air-conditioning units is now less than the 1954 values quoted in the paper for 3-hp and 5-hp sizes. Recent surveys made in our territory show that the cost differential has dropped very appreciably during the last 2 years. We sincerely hope that this trend continues.

The authors have shown conclusively that single-phase utilization in residential areas provides investment savings to the utility. Our company is now completing a study of the total economics, i.e., investment and operating costs of the utility plant plus investment and operating costs of the customers equipment. (The customers eventually pay for all of these.)

We recognize that the problem of holding the line on expansion of 3-phase utilization in residential areas has many facets, not the least of which is maintenance of good relations with customers, dealers, and housing developers. However the rewards are high, amounting to many millions of dollars, and we recommend that the conclusions of this paper be recognized and accepted; that we examine our company rules pertaining to extension of 3-phase service in residential areas and adjust them if necessary; that we examine our rules pertaining to maximum starting current of single-phase motors and adjust them, if necessary, to accommodate 5-ton air-conditioning installations; and that our customers and dealers be fully informed of the merits of single-phase service up to and including 5-ton units as well as the merits of multiple single-phase units where more than 5 tons are required.

N. H. Erlandson (Carolina Power and Light Company, Raleigh, N. C.): The authors have presented a very timely paper which deals with a problem confronting many utilities today, with the results indicating that the sum of the utility investment and the customer investment is the lowest for single-phase installations.

The indications are now that the future markets for home air conditioning are almost unlimited. At the present time, in the south at least, the single-phase room-type window units of 1 hp or less have taken over the market. This may be due to standardization of this type of unit, but its use is definitely affected by the cost of installation and in rate differentials applied by numerous utilities to cover the additional cost of serving small 3-phase loads.

There are other facts that favor single-phase installations that are difficult to evaluate but should be considered. In areas with no series street-light circuits, pole-top pin construction can be used for single-phase laterals. This type of construction requires less tree trimming, is subject to fewer outages, and is less noticeable than crossarm construction. Since operating personnel are accustomed to single-phase installation in residential areas, the installation of 3-phase units would result in both single-phase and 3-phase services, meters, and equipment in residential areas which will make installation and operating problems more difficult. Future conversions to higher voltages would also be more costly with 3 phase than single phase.

In making a study of this kind, certain assumptions must be made; however, there is one point that is not clear and that is the reason for the secondary conductor sizes for existing systems used in the paper. Tables III and IV for 3- and 5-hp air-conditioning units are based on an existing system consisting of 25-kva transformers, six spans of no. *4* copper secondary, four customers per pole, and load per customer of 1.25 kva. Tables V and VI for 3- and 5-hp heat pumps are based on existing system with the same conditions except the secondary size is no. *2* copper. If the existing secondary for the heat pump was no. *4* copper, it may be possible to stay within voltage limits and use the no. *4* copper to serve the 3-phase units while the single-phase units would require changing conductor. Since all other existing conditions were equal, it seems that a more realistic comparison would have been made by using the same size secondary conductor.

The 1.25-kva diversified demand per customer probably is applicable to areas with high saturation of gas appliances, but areas with high saturation of electric ranges and water heaters will have a much higher demand. A recent test on the Carolina Power and Light Company system showed a 3.6-kw peak diversified demand per customer at the transformer, with a 100% range and water-heater saturation but no air conditioning. A comparison of Tables VII, X, XI, and XII for 100% saturation indicates that higher lighting and appliance demands would increase the savings of single phase over 3 phase. It seems reasonable to conclude that an increase in the 1.25-kva demand figure used would increase the single-phase advantage in any of the cases studied.

A. H. Kidder (Philadelphia Electric Company, Philadelphia, Pa.): Review of this paper leads me to confirm the authors' conclusion that it is more economical from the power companies' standpoint to provide single-phase supply for residential air-conditioning loads. A spot check of the differential cost for the case of 4% saturation in Table III, indicates that the savings would be considerably lower but nevertheless appreciable on our system. The downward trend of savings with increased saturation is, of course to be expected, because the 3-phase cost would be relatively high initially in a single-phase area and would decrease naturally as the number of 3-phase loads per transformer installation increases.

A. S. Anderson and **Chase Hutchinson**: The number and length of discussions are such that we will attempt to reply only to the more pertinent issues that have been raised. We appreciate the interest in this subject and are glad to get conformation of views of the tremendous economic importance of making the proper decision as to whether the future air-conditioned home should receive single-phase or 3-phase service.

Mr. Lawrence's method of determining the unit savings is satisfactory provided there is a good way of allocating the extra cost of 3-phase units to the customers not using air conditioning.

Mr. Diehl states that experience to date indicates that the polyphase unit is more rugged than a single-phase unit. Other manufacturers using hermetically sealed units have not made this observation. There is need to collect and analyze data which are available by some central agency. His remarks emphasize the need for protection of all three phases of 3-phase units.

Mr. McClure asks if the same conclusions hold regardless of primary distribution voltage. The answer is yes. This may be seen from the tables. Note that the savings in primary decreased to an insignificant amount for a comparatively low saturation. With lower primary voltage the decrease would occur at somewhat lower saturation but the net result would remain the same.

The discussion by Mr. Ambrosius and Mr. Sarikas emphasizes the need for complete 3-phase protection which is lacking on all units except of one manufacture. Their viewpoint with regard to probability of more outages with a 3-phase unit is a practical consideration.

Mr. Johnson's plea that the industry needs to study the flicker problem has been voiced by the authors repeatedly before the AIEE and the Edison Electric Institute. We agree most wholeheartedly with Mr. Johnson that modern experience shows the customers will tolerate much more flicker than is presently allowed in published data on the subject. Secondary designs based on such data can be very costly.

Mr. McGrath asked some questions which are answered as follows:

1. Offhand, it would appear that the same-size single-phase units for residential service should be applicable to commercial service.

2. Our cost differentials of $50 compared to Mr. McGrath's quotation of $65 for 3-hp units and $80 compared to his $115 do not seem too far apart. We were indicating possible cost differentials if a substantial part of the market were all one phase or all 3 phase. He is quoting today's cost.

3. The extra cost of transformers and meters for 3-phase application is intrinsic.

4. We agree with Mr. McGrath's observation regarding the wiring costs. However, our estimates for a typical installation checked the average. This was coincidental, of course, but indicated the differential either way would not affect the final conclusion.

5. Some manufacturers still have the philosophy that 3-phase voltage should be balanced at the service switch. We are not in accord with this viewpoint and have urged both the AIEE and the Edison Electric Institute by formal papers of this kind to consider an industry study of the problem.

6. Locked rotor currents are about the same and locked rotor torque for single-phase motor is somewhat better. Also for the same investment, larger conductors can be installed for single-phase service. Therefore, the flicker for the single-phase system should be less. Actually, the study is made on an equal flicker basis wherever flicker is limiting.

Mr. Gallaher's company serves central-unit residential air conditioners with either single-phase or 3-phase motors. It is important, therefore, to note that Mr. Gallaher sees considerable operating advantages as well as investment savings in single-phase units.

Mr. Campbell states that the differential cost between a single-phase and 3-phase heat pump has been reduced to about $35 to the customer. We agree with his pithy conclusion: "Seldom do we get an answer to an economic analysis as clearcut as this one."

Mr. Brady and Mr. Mohr state the savings to the utility and eventually to the customer will run into millions of dollars. We wholeheartedly recommend that utilities take the steps recommended in their discussion.

Mr. Erlandson questions the use of different-size secondary for Tables V and VI than in Tables III and IV. Since some companies have used no. *2* copper as standard size for a great number of years, they would be interested in the case where existing copper of no. *2* size was used hence Tables V and VI. There is not enough difference in starting current of single-phase and 3-phase motors of same horsepower rating to make these two sizes of conductor equivalent from a flicker standpoint which is usually limiting.

The Electric System of the Rectifier-Type Locomotives for the Virginian Railway

J. C. BROWN
ASSOCIATE MEMBER AIEE

IN OCTOBER 1956 delivery began on an order of twelve 3,300-hp (horsepower) rectifier-type locomotives for The Virginian Railway Company. These locomotives mark the climax of 45 years of development since the first experimental mercury-arc rectifier locomotive was built, and represent the acceptance of this type of motive power for main-line railroad freight service.

While basically similar in characteristics and performance to the 4,000-hp locomotives delivered in 1955 to the New York, New Haven and Hartford Railroad for passenger service,[1] the present units are designed for low speed, heavy freight service on the 134-mile electrified zone of the Virginian Railway between Mullens, W. Va., and Roanoke, Va. In this service they will supplement the four 6,800-hp motor-generator locomotives acquired by the railroad in 1948, and will replace the original phase-converter locomotives that have been operating since the beginning of the electrification in 1926.

General Considerations

The performance of this locomotive in terms of tractive effort and speed is shown in Fig. 1. Selective auxiliary notching is provided on any of the 17 basic controller steps so that a maximum number of tractive-effort increments are obtained for optimum performance. This minimizes the possibility of slipping the wheels as the master controller is advanced, and assures extremely smooth starting of heavy tonnage trains.

The dynamic braking characteristic is shown in Fig. 2. The circuits are designed to give a peak braking effort of 63,000 pounds at 25 mph (miles per hour).

Description

The locomotive, classified as a road-switcher type, is shown in Fig. 3. It has a total weight of 396,000 pounds, all on drivers, and is $69^1/_2$ feet long over couplers. Location of the principal pieces of apparatus is shown in Fig. 4. The control equipment is of the electropneumatic type and is designed for single-end multiple-unit operation. The propulsion equipment consists of six General Electric Company *752* motors geared for a maximum locomotive speed of 65 mph with 40-inch diameter wheels. The total weight of the locomotive is supported on two 3-axle rigid bolster trucks through a centerplate and loading-pad arrangement.

One spring-raised air-lowered pantograph is located over the short-end cab at the front of the locomotive to collect current from the 11,000-volt 25-cycle trolley wire. A high-voltage bus line is provided for multiple-unit operation. The main transformer is Pyranol*-filled and force-ventilated. The rectifier equipment, consisting of 12 ignitron rectifier tubes and their associated firing circuits, is housed in two cubicles. The auxiliary rotating equipment includes one motor-generator set for 75-volt control power and battery charging, one motor-driven axial-flow equipment blower, one motor-driven air compressor, a water pump for rectifier cooling, and two axial-flow double-end resistor blowers. All the electric equipment is arranged to give ready access for inspection and maintenance.

Electric Apparatus

The prime objective in the development of the electric system for this locomotive is to produce a reliable, low-cost

---

Paper 57-32, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21-25, 1957. Manuscript submitted October 22, 1956; made available for printing November 27, 1956.

J. C. Brown is with the General Electric Company, Erie, Pa.

*Registered trade-mark of General Electric Company.

Vol. 7
Pt. 2
1957

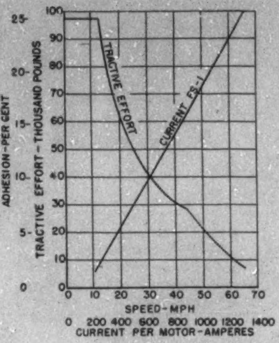


Fig. 1. Locomotive speed tractive-effort characteristic

system embodying the latest improvements in materials and processes together with any other factors that would contribute to long life and high availability. While the rectifier principle as applied to locomotive service is relatively new, the apparatus is new only in this combination and for this purpose. Each component has been proved through long service. The power-supply circuits are shown in Fig. 5, and the propulsion circuits in Fig. 6.

MAIN TRANSFORMER

The main transformer is rated 4,300 kva, 11,000/2,158/392 volts, 25 cycles, single phase, and has a center-tapped auxiliary winding. A 235-kva auxiliary notching transformer is included in the main transformer tank to save weight and space.

The transformer main secondary has taps at 2,158, 1,766, 1,439, 1,112, and 785 volts, which furnish a variable voltage supply to the rectifiers. The auxiliaries and buck-boost power are supplied from a center-tapped auxiliary secondary winding 196 volts on each side of the center tap. The main transformer assembly includes a high-voltage bushing transformer *CT1* and current transformers on the main secondary taps *X3*, *X4*, and *X5*, and on the auxiliary secondary taps *Y2* and *Y3*, for transformer protective relays.

The transformer primary is protected by thermal relay (*TPR*) which is energized from a tap on the high-voltage bushing current transformer. This relay operates to protect the transformer if high primary currents persist.

The transformer secondary taps are protected by instantaneous relays (*SPR*), *SPR1* to *SPR5* inclusive, energized from current transformers mounted on the taps inside the transformer tank. These current transformers and their associated tripping coils are arranged to give tap-to-tap protection by instantaneous tripping of the relays on all currents in excess of the normal surges encountered during tap changing.

Operation of the *TPR* and *SPR* relays will energize the ground switch (*GS*), thus grounding the pantograph. The flow of current through the ground switch partially trips the ground-switch relay (*GSR*) and isolates the pantograph so that it cannot be lowered remotely. When the track circuit breaker opens, it cuts off power to the trolley circuit of the locomotive and de-energizes the *GSR* relay. This then completes its cycle and energizes the pantograph down-magnet valve, lowering the pantograph.

The *TPR* and *SPR* relays are in the relay compartment in the operator's cab. A target is provided on each relay to facilitate locating a fault. All tap-changing equipment for the transformer is mounted on the side; see Fig. 7.

RECTIFIER EQUIPMENT

The rectifier equipment consists of twelve 12-inch *GL-6878* ignitron rectifier tubes, as shown in Fig. 8. These tubes are similar to those used on the locomotives for the New York, New Haven and Hartford Railroad, except that they are larger in diameter and each group of three tubes is connected to a common cathode. Four current-dividing reactors are provided to balance the current between individual tubes in each group.

The duration and magnitude of fault current resulting from an arc back in a main rectifier tube is limited by a flux-shifting type of circuit breaker. Its action is dependent on the arc-back sensing circuit producing the trip signal for the breaker *JR*, when one or more tubes conduct in the reverse direction.

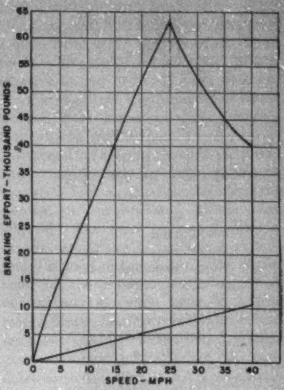


Fig. 2. Locomotive characteristics in dynamic braking

The basic elements of the arc-back sensing circuit, shown in Fig. 9, are the current transformers, *CT21A* to *21D* inclusive. These arc-back current-trip transformers have secondary windings through which a d-c saturating current flows, and 1-turn primaries through which the current of each group of three tubes passes. The current saturates the *CT* so that the normal direction of tube current can produce little flux change. The arc-back current opposes the saturating current ampere-turns, and tends to reduce the core flux to zero and drive it to saturation in the opposite direction. This large flux change, resulting from arc back, produces a voltage in the *CT* secondary which is applied to the trip coil of the *JR* breaker. Thus, the saturated current transformer discriminates between forward and reverse tube currents and provides a voltage pulse when arc back occurs. A relay *JROL* (Fig. 5) is pro-

Fig. 3. Three-quarter view of completed locomotive

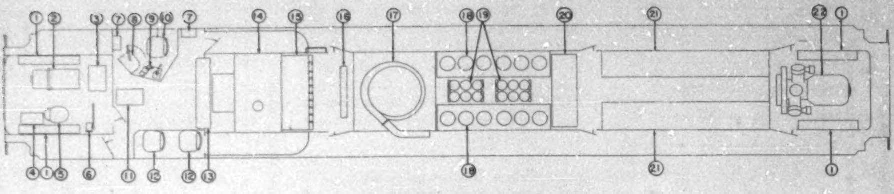

1
2
3
7
8
9
10
7
14
15
16
17
18
19
20
21
22
1
4
1
5
6
11
12
12
13
18
21
1

Fig. 4. Location of principal pieces of apparatus

1—Sand box
2—Motor-generator set
3—Air-brake equipment
4—Clothes locker
5—Toilet
6—Hand brake
7—Relay cabinet
8—Master controller
9—Air-brake control stand
10—Engineman's seat
11—Water cooler
12—Crew seat
13—Relay compartment
14—Transformer
15—Tap switches
16—High-speed breaker
17—Equipment blower
18—Rectifier tubes
19—Reactors
20—Rectifier cooling equipment
21—Control compartment
22—Air compressor

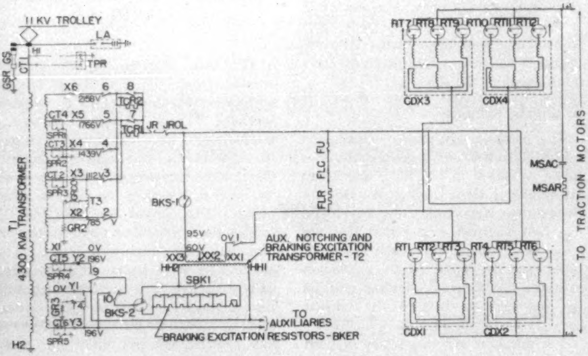

11 KV TROLLEY
LA
H1
TPR
GS
CT1
GSR
X6
6
8
258V
TCR2
7
CT4
X5
5
1766V
TCR1
JR
JROL
SPR1
CT3
X4
4
1439V
SPR2
CT2
X3
1112V
3
SPR3
T3
X2
2
785
GR2
4300 KVA TRANSFORMER
T1
X1
0V
CT5
Y2
196V
SPR4
9
0V
Y1
10
GR3
T4
BKS-2
CT6
Y3
196V
SPR5
H2
BKS-1
95V
60V
0V
XX3
XX2
XX1
HH2
HH1
SBK1
AUX. NOTCHING AND BRAKING EXCITATION TRANSFORMER - T2
TO AUXILIARIES
BRAKING EXCITATION RESISTORS - BKER
FLR
FLC
FU
RT7
RT8
RT9
RT10
RT11
RT12
CDX3
CDX4
MSAC
MSAR
RT1
RT2
RT3
RT4
RT5
RT6
CDX1
CDX2
TO TRACTION MOTORS

Fig. 5. Locomotive power-supply circuits

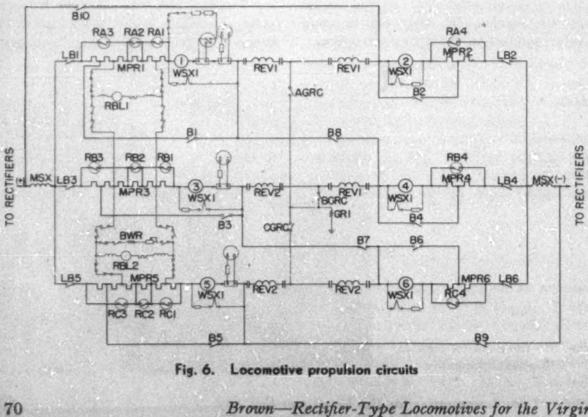

B10
RA3
RA2
RA1
LB1
MPR1
1
WSX1
REV1
REV1
2
WSX1
RA4
MPR2
LB2
B2
AGRC
RBL1
B1
B8
RB3
RB2
RB1
RB4
MSX
LB3
MPR3
3
WSX1
REV2
REV1
BGRC
GR1
4
WSX1
MPR4
LB4
MSX(−)
B3
B4
CGRC
BWR
RBL2
B7
B6
LB5
MPR5
5
WSX1
REV2
REV2
6
WSX1
MPR6
LB6
RC4
RC3
RC2
RC1
B5
B9
TO RECTIFIERS
TO RECTIFIERS

Fig. 6. Locomotive propulsion circuits

vided to trip the *JR* breaker on sustained forward fault.

The magnitude of the arc-back current in the ignitron is also dependent on the a-c system reactance. The current magnitude will therefore depend upon the location of the locomotive on the power system, and will vary over a wide range. The various levels of arc-back current tend to produce wide variations in the *CT* secondary-voltage amplitude and duration. A Thyrite* resistor load *R201* is placed across the secondary of each current transformer. This resistor exerts a nonlinear loading effect on the *CT* secondary, thereby increasing the pulse duration and limiting the crest magnitude. The loading effect is more pronounced during high-current arc backs which tend to produce short-duration high-voltage pulses. Thus, this Thyrite tends to make the secondary-voltage pulse of constant amplitude and duration.

A counter *ABC* is associated with each current transformer. This counter is operated from the *CT* secondary voltage obtained during arc back. The *CT* secondary voltage charges a capacitor *C505* through a resistor *R202*. The capacitor then discharges through the counter coil and yields a current pulse of sufficient duration to operate the arc-back counters. These counters give an accurate record of the number of arc backs for each group of three tubes.

Each rectifier tube has two ignitors, one is pulsed and the other is a spare. The function of this ignitor is to initiate a cathode spot on the mercury-pool cathode. This function must be per-

---

*Registered trade-mark of General Electric Company.

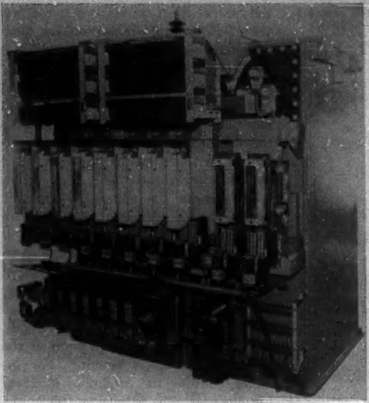


Fig. 7. Transformer with associated switching equipment mounted on case

formed at the proper time in relation to the anode voltage and must be repeated once each cycle. The ignitor is energized by a short-duration current pulse which is supplied by the ignitor firing circuit.

A sine-wave voltage is applied from the 350-volt rectifier excitation transformer to a nonlinear resonating circuit. This circuit produces two firing pulses of approximately 30-degree width and opposite polarity. These pulses are spaced 180 degrees apart and are transmitted to the ignitor through the ignitor coupling transformer.

The rectifier tubes have two additional grid electrodes provided for reliable operation at rated load. Next to the anode is an inner or control grid which is a perforated cage surrounding the anode. Surrounding the control grid, and between it and the cathode, is another grid of similar construction. This is the shield or outer grid.

A special grid circuit supplies the correct potential for firing the grids. This circuit incorporates a permissive feature which prevents the application of positive voltage to the inner grid unless a cathode spot has been started. If no cathode spot exists, the outer grid cannot conduct and no positive voltage will be applied to the inner grid.

A neon lamp is supplied with each rectifier tube to indicate misfires or failure of the ignitor to start a cathode spot. The lamps normally light on each grid-firing pulse with a flicker rate of 25 cycles. During a misfire, the outer grid will not conduct current, therefore, the neon lamp will fail to light. Misfires of 1-cycle duration can be readily detected by observation of the lamp.

Ideal operating conditions for an ignitron rectifier require that the water in the cooling jacket be maintained at fairly constant temperature. For this reason, temperature-regulating equipment is provided. Under normal conditions, water temperature at the inlet to the main tubes is between 35 C (degrees centigrade), which is 95 F (degrees Fahrenheit) and 50 C (122 F). Outlet temperature is approximately 43 C (110 F). A high-water-temperature alarm switch is supplied to flash an indicating light on the engineman's control panel if the temperature exceeds 62 C (140 F).

The coolant consists of approximately 75 gallons of distilled water containing 0.1-per-cent anhydrous sodium chromate as a corrosion inhibitor. Methanol alcohol is used to protect against freezing.

Each rectifier-tube water jacket has an electric potential difference to the rectifier frame, water pump, heater tank, and hose headers. The cooling solution in the system has sufficient conductivity to pass an appreciable current along the water stream through each hose. Sufficient length of hose is used to limit the magnitude of this current to safe values. To prevent electrolysis of the disconnect couplers at the tube end of each hose, a target is mounted in each hose nipple so that current enters and leaves the metallic part of the system at the target, rather than through parts of the conductor.

TRACTION MOTORS AND CONTROL

The traction motors are permanently connected, two groups in series and three groups in parallel; see Fig. 6. Variable voltage steps are provided by a combination of tap changing, auxiliary notching control, and resistance acceleration. The tap changing and resistance notching give a basic total of 17 tractive-effort increments. The auxiliary notching control is selective on any of the 17 basic notches. Thus, a relatively large number of small tractive-effort increments are provided for optimum performance.

In order to eliminate much of the interlocking and the large number of contactors required for a resistance-acceleration scheme, a cam-operated switch was selected. The cam switch, shown in Fig. 10, consists of 12 individual contactors, 4 in each motor circuit. The ohmic value of the resistances short-circuited out by the cam switch varies as a geometric progression; i.e., the value of the resistance short-circuited out by each cam contactor is half that of each other section. Thus, in a 4-section resistor, there are 16 possible steps. By selecting the maximum value of resistance for step 16, as required for the minimum starting current, the 16 steps can provide 9 desired steps with ohmic value sufficiently near the theoretical required values to be acceptable. Although individually operated contactors could be used, the transition between notches would not be as smooth as that obtained with the accurate timing of contactors actuated by a

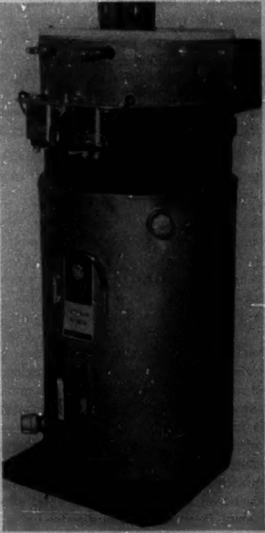


Fig. 8. Locomotive-type 12-inch rectifier tube

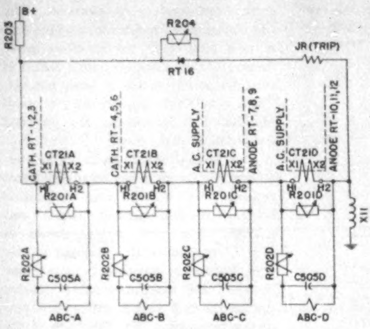


Fig. 9. Arc-back sensing circuit

common cam shaft. The rather high inductance of the smoothing reactor will tend to maintain the current in the circuit at its initial value while the resistance changes are being made and, therefore, there will be little or no torque fluctuation on the traction motors.

The accelerating resistor is composed of three identical sections. These sections consist of ribbon-type resistor units blown with two double-end axial-flow fans. For normal operation, blower *RBL1* (Fig. 6) is connected across the resistor section in motor circuits 1 and 2. The second blower *RBL2* is connected across the resistor section in motor circuits 5 and 6. A switch is provided for cutting out the motors by pairs. If motor 1 and 2 are cut out, blower *RBL1* is transferred to motor circuits 3 and 4. If motors 5 and 6 are cut out, blower *RBL2* is transferred to motor circuits 3 and 4.

Wheel-Slip Protection

Wheel-slip protection is provided to detect any slipping axle or any two simultaneously slipping axles in motor groups 1, 3, and 5; or 2, 4, and 6. Conventional wheel-slip equipment which uses relays does not work satisfactorily when applied to traction motors operating on rectified alternating current. Conventional d-c relays require the use of elaborate filters because of the presence of ripple current in the circuit. To overcome these difficulties, a saturable reactor is used to compare the motor voltages or currents. This is done by arranging control windings so that when all motor quantities are balanced, the net control excitation is zero. When a wheel slip occurs, these motor quantities are unbalanced and excitation results in the control reactor. This initiates corrective action. A second-stage magnetic amplifier is used to improve the sensitivity of the system.

The reactor *WSX1*, shown in Fig. 11, is a series-connected saturable reactor with control windings connected to read motor armature voltage. If the voltages across all motors are balanced, the net ampere-turns produced in the six windings are zero, and there is substantially no output from reactor *WSX1*. As an example of the operation of the system, consider the no. 5 motor circuit. Reactor winding *F5-F6* is connected, as shown in Fig. 6, reading motor voltage on the armature of motor 5. When a wheel slip occurs on this motor, the generated voltage in the armature increases. With this increase in voltage across motor 5 there is a corresponding decrease in voltage across motors 2, 4, and 6, since the total output voltage from the reactor is substantially constant. Thus, the current is increased through the winding *F5-F6* and, to a lesser extent, through the *F1-F2* and *F3-F4* windings. At the same time, the current through windings *F11-F12*, *F9-F10*, and *F7-F8* is decreased. This unbalance in current between the six control windings increases the current flow through reactor *WSX1*, since the total ampere-turns in the *A1-A2* winding must match the total ampere-turns produced by the control windings, a characteristic of the series saturable reactor. The current from *WSX1* passes through the rectifier *RT21*, and is used to energize the *F1-F2* control winding on the magnetic amplifier *WSX2*. This amplifier is normally biased to an off position by winding *F3-F4* which receives its current from a battery signal. However, current flow through the winding *F1-F2* overcomes this battery signal and increases the output of *WSX2* until it becomes great enough to pick up the relay *WS1*. This relay operates the wheel-slip light and buzzer, and energizes a time-delay relay to cut off power if corrective action is not taken. The sensitivity of the overall system is determined by the amount of battery signal in the *F3-F4* winding.

The effect of parallel motors is also obtained by a power tie, shown in Fig. 6, connecting the mid-voltage point of the three 2-motor series circuits. For instance, consider the conditions if motor 1 is slipping. The voltage on this motor will rise and, because of the power tie, the voltage on motors 3 and 5 will also rise. There will be a corresponding decrease in the voltage on motors 2, 4, and 6. The current in motor 1 decreases because of its slipping. The current in motors 3 and 5 increases, and that in motors 2, 4, and 6 decreases. Part of the current from motors 3 and 5 then flows through the power tie to motor 2. As a result, there is a voltage build-up on non-slipping motors, thus minimizing the tendency of the slipping axle to overspeed, and making it easier to recover from a slip. Also, one of the heavily loaded axles receives the current build-up, and the more lightly loaded axles receive a current reduction; current reduction is minimized in the motor which is paired with the slipping motor. The net result is that tractive effort is maintained at a high value even though an axle slips.

Ground Protection

The ground-relay system consists of three parts which provide protection for the traction motors, the rectifier, and the transformer. The system is selective, and restricted locomotive operation is available if only one ground exists. The motor ground relay *GR1* operates a light and bell signal. When this relay picks

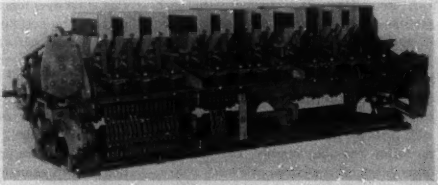


Fig. 10. Cam-operated accelerating switch

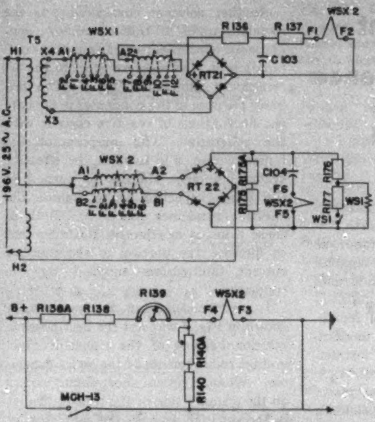

WSX 2
R 136
R 137
F1
F2
WSX 1
T5
H1
X4
A1
A2
C 103
RT21
X3
196 V. 25 ~ A.C.
WSX 2
A1
A2
RT 22
R175A
R175
C104
F6
WSX2
F5
R176
R177
WS1
WS1
B2
B1
H2
B+
R138A
R138
R139
WSX2
F4
F3
R140A
R140
MGH-13

Fig. 11. Wheel-slip detection and correction circuits

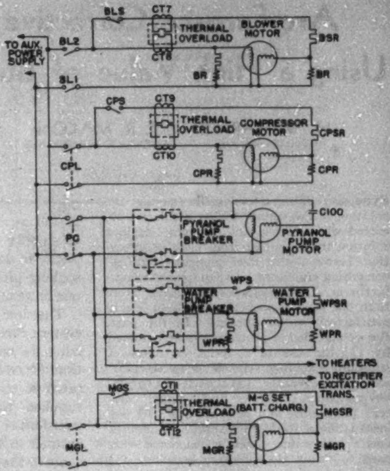

TO AUX. POWER SUPPLY
BLS
CT7
THERMAL OVERLOAD
BLOWER MOTOR
BSR
BL2
CT8
BR
BR
BL1
CPS
CT9
THERMAL OVERLOAD
COMPRESSOR MOTOR
CPSR
CT10
CPL
CPR
CPR
C100
PYRANOL PUMP BREAKER
PYRANOL PUMP MOTOR
PC
WPS
WATER PUMP BREAKER
WATER PUMP MOTOR
WPSR
WPR
WPR
TO HEATERS
TO RECTIFIER EXCITATION TRANS.
MGS
CT11
THERMAL OVERLOAD
M-G SET (BATT. CHARG.)
MGSR
CT12
MGL
MGR
MGR

Fig. 12 (right). Schematic diagram of auxiliary circuits

up, it de-energizes the rectifier excitation contactor (*REC*), tap switches, and motor line contactors. The relay may be reset remotely, if the ground is cleared by cutting out the traction motors. If a ground occurs in the transformer or rectifier circuits, the *GR2* relay is energized. If the ground is in the rectifier circuit, restricted operation to the 11th controller notch is maintained. If the ground is in the transformer circuit, the relay must be manually reset. Indicating lights are provided to indicate the location of ground.

Dynamic Braking

During braking, each traction motor armature is shunted with a resistor. All six motor fields are connected in series across the rectifier output. The rectifier is excited from the 96-volt tap, *XX1-XX3* in Fig. 5, on the auxiliary notching transformer by connecting the latter to the low-voltage auxiliary winding.

Braking effort and speed are controlled by varying the voltage applied to the rectifier. This is done manually through the braking handle on the master controller which operates seven contactors, *SKB1* to *SKB7* inclusive, in the auxiliary notching transformer primary circuit to give 17 braking steps. A brake overload warning relay (*BWR*) is provided which lights the wheel-slip indicating light. Wheel-slip protection is also provided during dynamic braking.

Auxiliary Control

The various auxiliaries, shown in Fig. 12, are driven by single-phase 25-cycle induction motors connected to the 196-volt *Y1-Y3* tap on the auxiliary secondary winding of the main transformer. Split-phase resistance starting is used for these motors. A resistor is connected in series with the starting winding until the motor has reached approximately full speed. A balanced-beam type of relay then cuts out the starting winding. This relay has a restraint coil energized by the motor-line voltage, and a lift coil energized by voltage across starting winding. Thus, it compensates for trolley voltage changes and permits starting over a wide range of applied voltage.

All auxiliary motors except the transformer cooling pump and water-pump motors are protected against locked rotors by thermal trip relays with thermal elements in each winding circuit. The same protection is furnished the two pump motors by means of circuit breakers. The motor-generator set starting and line contactors are energized directly from the auxiliary secondary winding, thus making the starting of the set independent of battery condition. The starting and line contactors for the blower, compressor, transformer cooling pump, and water-pump motors are excited from the battery.

A relief valve unloads the compressor until it is up to speed. The compressor control is arranged so that the compressor runs continuously, and is loaded and unloaded by the compressor governor.

Summary

Experience gained in previous applications of rectifiers to locomotives has proved valuable in the design of this equipment. While these locomotives will operate under radically different conditions, test results have been highly satisfactory. They indicate that another forward step has been taken in the development of the rectifier locomotive as a valuable new addition to the motive power equipment available for electrified railroads.

Reference

1. RECTIFIER LOCOMOTIVES FOR THE NEW YORK, NEW HAVEN AND HARTFORD RAILROAD, F. D. Gowans. *AIEE Transactions*, vol. 74, pt. II, July 1955, pp. 183–89.

Arc-Furnace Corrective Equipment Using a High Value of Buffer Reactance

H. W. HARPER
NONMEMBER AIEE

T. R. MACON
ASSOCIATE MEMBER AIEE

A. F. SEDGWICK
MEMBER AIEE

Synopsis: This paper describes an effective solution to the voltage-flicker problem associated with electric arc furnaces. It is a solution that resulted from close co-operation between the steel company, utility, consulting engineers, and equipment manufacturers. The performance during more than one year of operation has been highly satisfactory in all respects. In the past, the voltage-flicker problem has been solved by (1) the use of a series capacitor, a standard reactance synchronous condenser, and a buffer reactor; (2) the use of a special low-reactance condenser and a buffer reactor; (3) the use of a separate bus. In this installation a standard reactance condenser and a high value of buffer reactance were used to obtain the following benefits: (1) satisfactory flicker performance, (2) power-factor correction, (3) automatic voltage regulation, (4) low installed cost, and (5) low maintenance cost.

SINCE the earliest installations of electric arc furnaces, engineers have been faced with the problems of the objectionable voltage flicker associated with furnaces. This voltage flicker has resulted from the rapid, repeated fluctuations of current in the furnace arc. The magnitude of the current fluctuations is directly related to the size and design of the furnace and associated equipment. The effect of the current fluctuations, i.e., voltage flicker, is related to the parameters of, and the configuration of, the power-supply system and corrective equipment. These are some of the problems which faced Northeastern Steel Corporation and the United Illuminating Company in the installation described in this paper.

Northeastern Steel was incorporated in Bridgeport, Conn., in 1954. It acquired from the Stanley Works its properties located in Seaview Avenue, Bridgeport, including principal production units as follows: three basic open-hearth furnaces with a combined annual ingot capacity of 188,230 net tons, two cupolas, soaking pits, one 34-inch blooming mill, one billet mill, and two hot strip mills.

The first step in the plans of Northeastern Steel in expanding and modernizing its facilities included the installation of two electric arc furnaces, each rated 35 tons. Each furnace has power supplied through a 10,000/12,500-kva transformer. The transformer primary is rated 13,800 volts and is connected by cable to the East Main Street Substation of the United Illuminating Company.

The United Illuminating Company has a generating capacity of 336,000 kw. The East Main Street Substation is the nearest source available to supply the load at Northeastern Steel. This is a 13.8-kv substation with a double bus. The available short circuit at East Main Street is approximately 800,000 kva. The East Main Street Substation supplies approximately 10,000 kva of residential load and 15,000 kva of other industrial load.

In order to eliminate objectionable voltage flicker, it was decided that the Northeastern Steel system should be designed to limit frequent voltage flicker to 0.5% at East Main Street. This value of flicker voltage was obtained from recognized voltage-flicker charts for frequent fluctuations (i.e., those fluctuations which occur several times per minute.) The 3-phase reactive fluctuations were therefore limited to 4,000 kilovars at the East Main Street Substation. It was estimated that the maximum frequent fluctuations of the furnace would be 8,000 kilovars, 3 phase.

Solutions Considered

It was decided that the connection of the furnace transformers directly to the East Main Street bus would not be satisfactory because of excessive voltage flicker. It was also decided that it would not be feasible to establish a separate bus at East Main Street to supply power to the furnace because of the expense and lack of operating flexibility.

Another solution considered was the installation of a buffer reactor, synchronous condenser, series capacitor combination at Northeastern Steel in the arrangement shown in Fig. 1. The primary purpose of the condenser is to share the fluctuations of reactive current with the generators. The purpose of the buffer reactor is to increase the effective system reactance whereas the purpose of the series capacitor is to reduce the effective condenser reactance. Both of these changes in effective reactance act to increase the portion of the reactive current fluctuations supplied by the condenser.

It was decided that the arrangement shown in Fig. 1 was not the best possible solution because of the transient overvoltage requirements of the series capacitor. When a system short circuit occurs on the primary side of the transformer, it is necessary to by-pass the capacitor by means of a gap to protect it from damage from overvoltage. In order to prevent the protective gap from flashing over on normal furnace fluctuations, it is necessary to include appreciable voltage margin in the capacitor. To have provided a series capacitor with sufficient voltage margin and with the necessary protective equipment would have increased the cost of this installation. In addition, maintenance costs would have been increased by that required for the capacitor protective equipment.

The use of a buffer reactor and special low-reactance condenser was also considered. This solution was not used because it would have resulted in increased installed cost.

Solution Used

Although a high value of buffer reactance has been considered[1] likely to contribute to condenser instability, a recent paper[2] suggested that the reactance can safely be made greater than had been previously used. After thorough analysis of all considerations, it was decided that the solution shown in Fig. 2 would meet the operating requirements. The value of buffer reactance was made high enough to obtain the proper division of reactive current without resorting to the use of a special low-reactance condenser or a series capacitor. This solution was found to be more economical than a comparable solution using a low-reactance condenser, or a standard reactance condenser with a series capacitor. In addition, it is a solution free of the complexity and operational problems associated with the series capacitor.

Paper 57-55, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21-25, 1957. Manuscript submitted October 22, 1956; made available for printing November 29, 1956.

H. W. Harper is with the Northeastern Steel Corporation, Bridgeport, Conn.; T. R. Macon is with the General Electric Company, New Haven, Conn.; and A. F. Sedgwick is with The United Illuminating Company, New Haven, Conn.

The authors wish to express their appreciation for the assistance provided by their associates. Jackson and Moreland acted as consultants for The United Illuminating Company.

Vol. 7
Pt. 2
1957

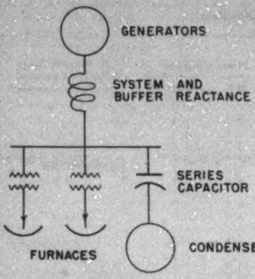
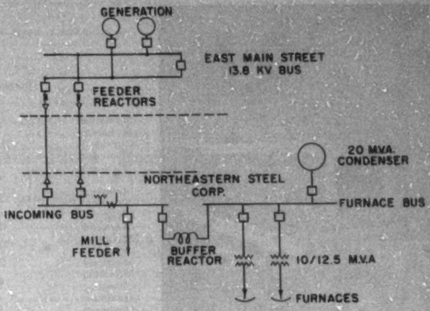


Fig. 1. One-line diagram of arc-furnace power system including a series capacitor

Fig. 2. One-line diagram of arc-furnace power system at Northeastern Steel Corporation

Per-unit reactances = 20-mva 13.8-kv base
Generation = 0.025
Feeders (each) = 0.070
Condenser $X_{d'}$ = 0.48, $X_{d''}$ = 0.28
Buffer reactor = 0.49
Furnace and transformer (each at maximum second-volts) = 0.62



Equipment Description

One of the furnaces is shown in Fig. 3. Each furnace is rated 35 tons with the normal charge being approximately 50 tons. The electrodes are positioned individually by d-c motor drives controlled by amplidyne generators. Each of the two furnace transformers is rated 10/12.5 mva (megavolt-amperes), 13.8 kv primary with a maximum of 285 volts secondary. Fig. 4 shows a view of one of the transformers before tanking.

The condenser, shown in Fig. 5, is a standard, air-cooled machine rated 20 mva, 13.8 kv, 720 rpm. The air coolers are located in a pit below the foundation. The exciter is direct connected. An amplidyne voltage regulator is used to control the excitation.

The switching equipment shown in part in Fig. 6 consists of metal-clad switchgear with magnetic air circuit breakers. The one-line diagram is shown in Fig. 2. The short-circuit duty on the incoming bus at Northeastern Steel is approximately 335 mva, and is expected to increase. Due to the high value of buffer reactance, the short-circuit duty on the furnace bus is limited to approximately 110 mva. Breakers which are rated 500 mva interrupting capacity are used throughout the switchgear to (1) provide interchangeability of all circuit breakers, (2) reduce the severity of the duty and thus frequency of inspection and maintenance of the furnace breakers, (3) make unnecessary the use of a separate breaker to back up the furnace breakers. The use of a backup breaker is customary when furnace breakers are applied near their interrupting capacity.

Another benefit provided by the high value of buffer reactance is that it makes across-the-line starting of the condenser feasible; thus neither a starting breaker and reactor nor a starting motor is necessary. On starting, the inrush is limited to approximately 25 mva at 13.8 kv.

Because of the high value of buffer reactance, and resulting low value of short-circuit current, two sets of voltage-restrained overcurrent relays are used in the metal-clad switchgear to trip the condenser and reactor breakers. The sum of the buffer reactor and condenser currents, along with the condenser terminal voltage, are used to operate one set of three relays. This set of relays is unaffected by condenser starting or condenser swinging and provides fast clearing for faults on the furnace bus, as well as providing backup for the furnace breakers. The other set of three voltage-restrained overcurrent relays, using reactor current and incoming bus potential, provide fast clearing of faults on the incoming bus.

A demand-limiting meter is used to limit the kilowatt demand to a predetermined level. As initially connected, the demand-limiting meter was used to (1) reduce the power level and (2) if after a time delay the demand had not been reduced below the ideal demand, trip the furnace breakers. That portion of the demand-limiting circuit which tripped the furnace breakers has been made inoperative because it hampered production and imposed unnecessary shock loading on the utility system. As now connected, the power level of each furnace is reduced approximately 2.5 megawatts if the ideal demand is exceeded.

The buffer reactance is provided by

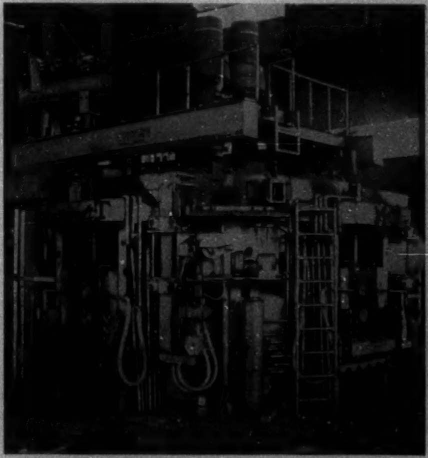


Fig. 3. 35-ton electric arc furnace

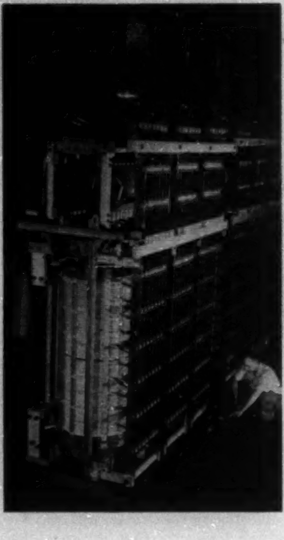


Fig. 4 (left). 10/12.5-mva arc-furnace transformer prior to tanking

Table I. Representative Data

	Furnace, Mvar	Condenser, Mvar	Power Company		% Furnace Mvar Change Furnished by Power Company
			Mw	Mvar	
Maximum	11.7	14.7	13.0	6.1	
Minimum	0.7	2.3	3.3	1.4	
Change	11.0	12.4	9.7	4.7	42.6
Maximum	13.4	16.7	13.6	4.7	
Minimum	3.8	6.3	5.2	1.7	
Change	9.6	10.4	8.4	3.0	31.2
Maximum	13.2	17.0	13.5	4.7	
Minimum	4.5	5.7	9.5	1.7	
Change	8.7	11.3	4.0	3.0	34.5
Maximum	13.4	17.0	13.3	4.0	
Minimum	5.0	6.6	8.1	1.3	
Change	8.4	10.4	5.2	2.7	32.1
Maximum	12.4	17.0	13.3	4.5	
Minimum	5.4	6.0	10.0	1.4	
Change	7.0	11.0	3.3	3.1	42.8
Maximum	6.7	5.5	3.5	4.6	
Minimum	0.0	−2.9	0.0	1.2	
Change	6.7	8.4	3.5	3.4	50.8
Maximum	7.9	10.5	11.5	3.7	
Minimum	1.9	2.0	5.0	1.2	
Change	6.0	8.5	6.5	2.5	41.7
Maximum	9.6	13.5	12.7	4.4	
Minimum	6.7	9.2	11.0	2.9	
Change	2.9	4.3	1.7	1.5	51.7

three single-phase cast-in-concrete reactors, each rated 745 amperes, 13.8 kv, 43% drop on a 17.8-mva base (48% on 20-mva base). The reactors are mounted in a nonmagnetic enclosure. These enclosures are seen in the background, Fig. 6.

Selection of Ratings

The synchronous condenser voltage rating of 13.8 kv was selected so as to be the same as the voltage rating of the furnace transformers and the nominal system voltage. The condenser voltage regulator is set to hold approximately 13.8 kv. At present, when the furnaces are not operating, the condenser floats on the line. As the furnace demand is increased, the condenser supplies the reactive demand of the furnaces and one-half the reactive losses incurred between the regulated power company bus and the furnace bus. This results in a high power-factor load on the utility system. The selection of the kilovolt-ampere rating of the condenser and the current rating of the buffer reactors were based on the operating conditions of one furnace melting and the other furnace refining with some margin for occasional overlap of the melting periods.

The guaranteed rated voltage transient reactance of the condenser is 48%. The buffer reactance was selected initially so that the condenser transient reactance would equal the sum of the buffer and system reactances. The condenser would then supply at least half of the frequent fluctuations of reactive. It was decided to hold the starting inrush to the condenser to 25 mva to limit the voltage drop to approximately 3% at East Main Street. To meet this limitation, the buffer reactance was increased to give a total reactance of 55% on a 20-mva base.

The total reactance of 55% meets the stability criterion as stated in reference 2: "A suggested criterion for stable oper-

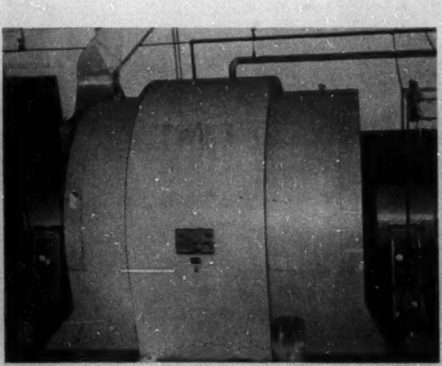


Fig. 5. 20-mva synchronous condenser

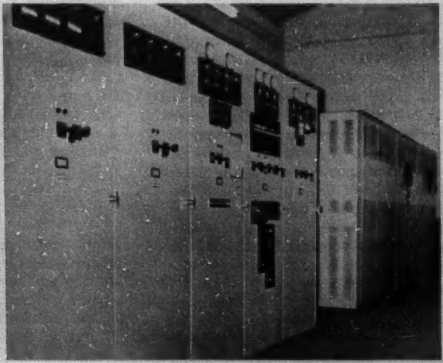


Fig. 6. 500-mva metal-clad switchgear with reactor enclosure in background

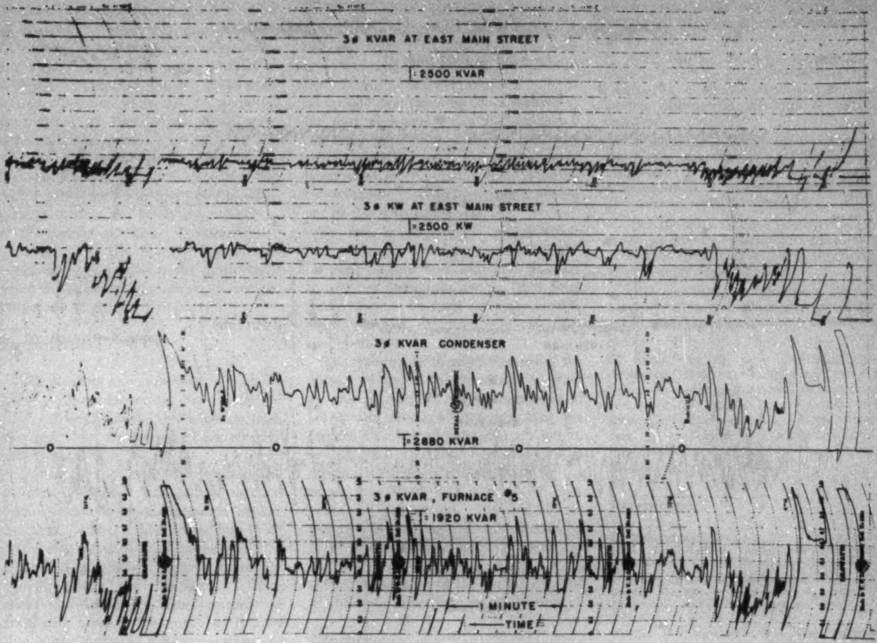


Fig. 7. Charts for one furnace melting

ation, as far as disturbances arising at the arc furnace are concerned, is that the buffer (and system) reactance be kept smaller than the effective reactance of the arc-furnace transformer and secondary leads. This criterion is seen to be independent of the condenser reactances." The study recorded[2] considered the case of supplying one furnace. It was realized that two furnaces would make the problem of condenser stability more severe; however, experience has shown[3] that the flicker caused by two furnaces was only slightly greater than that caused by one furnace. In addition, the operation of the furnaces was to be staggered so that two furnaces would not be melting at once.

Performance Tests

Tests were made with (1) one furnace melting, (2) simultaneous operation of two furnaces melting, (3) simultaneous operation of two furnaces with one melting and the other refining. Recording instruments with a chart speed of 3 inches per minute were used to record kilowatts and kilovars for: each furnace, the condenser, and the supply at East Main Street. Voltage at East Main Street was recorded. Charts were synchronized at the beginning and end of the tests by opening and closing the furnace breakers with a high current setting on the electrode regulators.

The condenser kilowatt charts were not satisfactory because the natural period of the condenser was so short that faithful instrument response was not obtained. Oscillograms obtained during subsequent tests, which show the kilowatt swings of the condenser, are included in reference 4. Tests were made with one furnace melting, and with two furnaces melting. The oscillograms obtained indicate an adequate stability margin.

Fig. 7 shows representative charts for a portion of a 10-minute test period with one furnace melting. A chart showing furnace power is not included since it is approximately the same as the power taken at East Main Street. The flat portions of the kilovar charts are due to the recorder going off scale. The test period started after 20 minutes of melting at reduced tap on a cold charge. The transformer taps and electrode regulators were set to obtain maximum furnace voltage and current. During this 10-minute test period, the fluctuations in 3-phase furnace reactive were: 25 fluctuations were 6 mvar (megavars) or greater, 9 were 8 mvar or greater, 4 were 10 mvar or greater, 3 were 11 mvar or greater. One of the three fluctuations greater than 11 mvar was caused by a cave-in; the other two were on the initial strikings of the arc. The most violent fluctuations occurred immediately after the initial striking of the arc when single phase arcing existed.

Representative data showing the maximum, minimum, and change in megawatts and megavars are shown in Table I. The per-cent furnace megavars change furnished by the power company was obtained by dividing the change in power company megavars by the change in furnace megavars. For larger fluctuations, the portion of the reactive furnished by the power company was, in general, less than

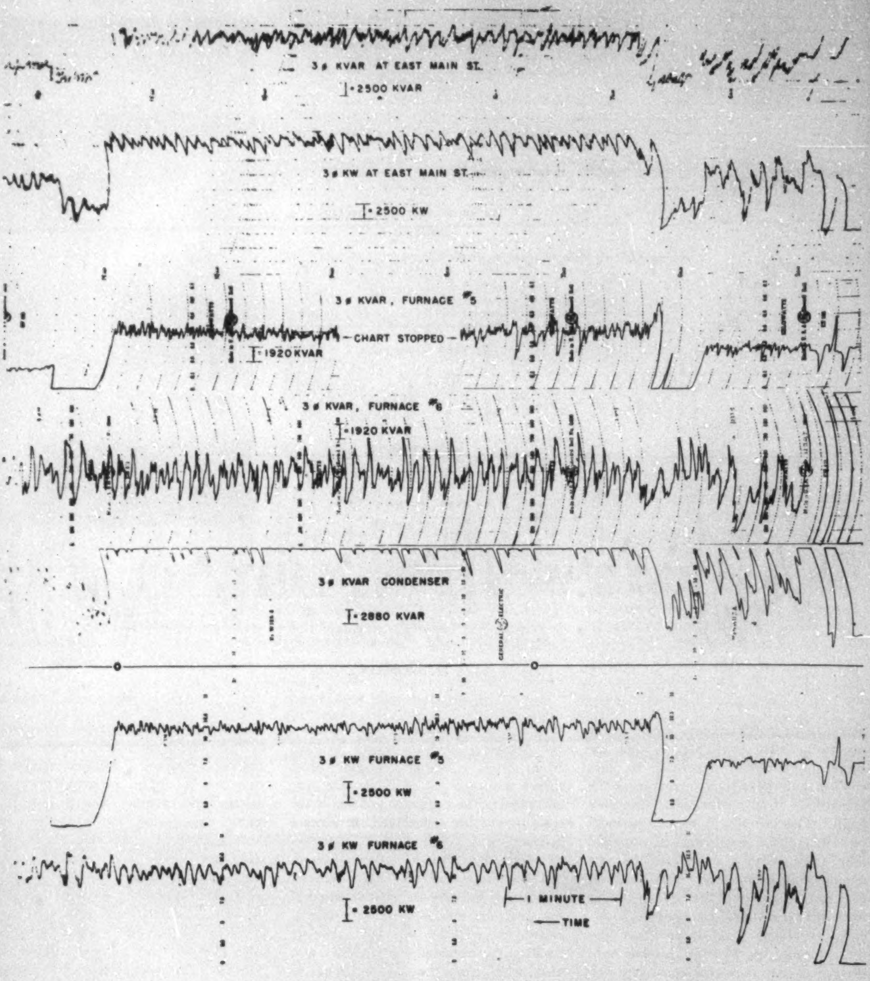


Fig. 8. Charts for two furnaces melting

that furnished on the smaller fluctuations. Several seconds were required for the larger fluctuations to be completed. During that interval, the forcing action of the amplidyne voltage regulator at maximum boost increased the condenser excitation so that the condenser furnished a large portion of the reactive fluctuations. The sum of the power company and condenser megavars exceed the furnace megavars by the I^2X losses in the reactances. The maximum change in voltage at East Main Street of 1.0% occurred during initial striking of the arc. All of the other changes due to the furnace were less than 1% and were apparently within normal changes in bus voltage. All of the recorded voltage changes which occurred at East Main Street during the tests were gradual and so infrequent that they were not in the category of flicker voltages.

The charts shown in Fig. 8 were made with two furnaces melting simultaneously. The data obtained are primarily of academic interest since the system was designed for staggered furnace operation.

The maximum change of 3-phase reactive at East Main Street was approximately 7.5 mvar. This change occurred on striking of the arc. After approximately 2 minutes of operation, the arc in the cold furnace became more stable and the fluctuations at East Main Street were reduced to 4–5 mvar.

It is of interest to note that as the furnace killowatt demand increases, the buffer plus system reactive losses increase rapidly because the losses are a function of the current squared. This can cause an increase in the reactive fluctuations supplied by the generators, particularly when the reactive fluctuations are accompanied by large power fluctuations.

Experience

The furnaces and associated equipment have now been in operation more than one year. After the initial shakedown period, the experience to date has been entirely satisfactory.

During the first weeks of operation, the condenser breaker was tripped several times by single-phase operation of the furnaces. These tripouts were due to an initial low setting of a current balance relay and the lack of familiarity of the operators with the equipments. Since these conditions have been corrected, the condenser breaker has not been tripped due to single-phase furnace operation. It is of interest to note that the condenser furnishes a large portion of the negative sequence currents caused by single-phase furnace operation.

In reference 1, it was reported that instability of the condenser had occurred in a similar installation with a comparably high reactance. To date, there has been no evidence of instability in this installation during operation or tests.[4] It is recognized that repeated shock loading at the natural frequency of the condenser can cause instability; however, with the well-adjusted electrode and voltage regulators used in this installation, repeated shock loading is so unlikely that it should not be the limiting criterion.

To date, no objectional voltage flicker has resulted from the operation of the furnaces. There have been no complaints from other customers served from East Main Street. In fact, no objectionable flicker has been observed by the substation operators.

The training, skill, and judgment of the furnace operator has considerable effect on the severity of the fluctuations of power and reactive. By operating the furnace for a few minutes after the initial, or back charge, on reduced voltage and current settings, the magnitude of the fluctuations can be reduced without penalizing production. By properly selecting and training their furnace operators, Northeastern Steel has, since the initial shakedown period, been able to reduce the severity of the fluctuations and at the same time increase production.

Conclusions

1. Careful co-ordination of the available experience, data, and equipment, along with close co-operation between the steel company, power company, consulting engineers, and equipment manufacturers, during the planning and design stages will contribute immensely to satisfactory arc-furnace installations.
2. With the well-adjusted, modern electrode and voltage regulators used in this installation, condenser stability is not the limiting criterion in the type of system described.
3. At high load levels, the reactive losses of the buffer plus system reactances may necessitate the use of a lower value of reactance than the maximum value dictated by condenser stability.
4. For the large reactive fluctuations, the condenser usually furnishes a larger portion because of the forcing action of the amplidyne voltage regulator.
5. The training and experience of the furnace operators can materially reduce the severity of the fluctuations with no sacrifice in production.
6. Experience in this installation has demonstrated that the use of a high value of buffer reactance and a standard reactance condenser will result in (*a*) satisfactory flicker performance, (*b*) power factor correction, (*c*) automatic voltage regulation, (*d*) low installed cost, and (*e*) low maintenance cost.

References

1. ARC-FURNACE LOADS ON LONG TRANSMISSION LINES, **T. G. LeClair.** *AIEE Transactions (Electrical Engineering)*, vol. 59, Apr. 1940, pp. 234–42.

2. VOLTAGE DIP AND SYNCHRONOUS-CONDENSER SWINGS CAUSED BY ARC-FURNACE LOADS, **C. Concordia.** *AIEE Transactions*, vol. 74, pt. III, Oct. 1955, pp. 951–56.

3. LARGE ELECTRIC-ARC FURNACES—PERFORMANCE AND POWER SUPPLY, **B. M. Jones, C. M. Stearns.** *Ibid.* (*Electrical Engineering*), vol. 60, July 1941, pp. 763–69.

4. SELECTION OF BUFFER REACTORS AND SYNCHRONOUS CONDENSERS ON POWER SYSTEMS SUPPLYING ARC FURNACE LOADS, **C. Concordia, L. G. Levoy, C. H. Thomas.** *AIEE Transactions*, vol. 76, pt. II, 1957 (*Paper no. 57–56*).

◆

Discussion

C. Concordia and **L. G. Levoy** (General Electric Company, Schenectady, N. Y.): We are especially pleased to hear of the good results obtained in the installation described. The imminence of this installation was one of our main reasons for undertaking, or at least accelerating the pace of, our theoretical study reported in reference 4 of the paper. We were pleased that our preliminary results could be of assistance in coming to a decision as to the values of reactance and condenser rating that would be used, and to see the selection justified by actual experience.

We should like to comment on two specific points of the paper.

First, we consider the values of kilovar fluctuation recorded on Fig. 7 to be a satisfactory check of the original estimate of 8,000 kilovars, 3 phase, which was used as a frequent fluctuation in the selection of this equipment. Every such check that is obtained gives us more confidence in our ability to select a proper typical shock for future cases.

Second, it is noted in the body of the paper that the principal reason that the power company furnished a relatively smaller percentage of the kilovar change for the largest changes was that these large fluctuations took some time to be completed. We believe it might be desirable to qualify the authors' conclusion 4 in order to bring out this point, as it might otherwise be misinterpreted. The regulator will be significant in affecting voltage flicker only for those fluctuations that either occur slowly or are sustained for a few seconds.

B. L. Lloyd (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The AIEE Working Group on Arc Furnaces has recently completed an extensive survey of existing arc-furnace installations[1] and resulting flicker conditions. The purpose was to establish power-system characteristics which can be used as a criterion for assuring satisfactory flicker performance when supplying arc-furnace loads. The most useful characteristic which expresses the ability of the system to supply these loads is the "mutual drop constant," designated X_{mn}. This quantity is defined as the voltage drop from the equivalent internal generator voltage to the critical lighting load when 1.0 per unit current (based on furnace transformer name-plate rating) at zero power factor is drawn by the furnace transformer. All other loads are neglected in determining this quantity. The use of 1.0 per unit current does not imply that this is the load swing which the furnace will experience; it is merely a convenient base upon which to express X_{mn}.

The Working Group report[1] indicates that for two 12.5-mva furnaces, values of X_{mn} less than 1.2% will result in satisfactory flicker performance, values greater than 1.9% will result in objectionable flicker, and intermediate values will be borderline.

Using the impedances in Fig. 2 of the paper, the value of X_{mn} with the condenser in service is 0.73% and without the condenser is 1.56%. This would indicate that during normal conditions, flicker performance should be entirely satisfactory. With the condenser out of service, the flicker at the East Main Street bus would appear to be borderline.

It would be interesting to know whether the furnaces have been operated without the condenser and whether the lamp flicker was considered to be in the broad categories of "borderline" or "very objectionable" during this condition.

REFERENCE

1. SURVEY OF ARC-FURNACE INSTALLATIONS ON POWER SYSTEMS AND RESULTING LAMP FLICKER, AIEE Subcommittee Report. *AIEE Transactions*, vol. 76, pt. III, 1957 (*Paper no. 57-9*).

H. L. Bassett (General Electric Company, Philadelphia, Pa.): This paper shows the successful result of extending into a new area of reactance values, etc., some principles previously used. The electrode regulators for the furnaces on this installation also include a new extension of a basic principle, namely, by providing higher electrode speeds than were previously used. Heretofore, the average electrode speed in large electric furnaces has been 3 to 4 feet per minute. The Northeastern Steel furnaces are geared for electrode speeds of $7^{1}/_{2}$ feet per minute while melting steel. While the authors' charts still show considerable power fluctuation, the number of fluctuations has been reduced by the ability of the electrodes to reach new operating positions more quickly.

Among the various noises made by an operating arc furnace, the crackle of an arc breaking and restriking is distinctive. Such noises are seldom heard on the furnaces at Northeastern Steel. Other high-electrode speed furnaces with which this writer is familiar are also generally free of arc-break noises, although they are frequent at some periods on low-electrode-speed furnaces in the same shops.

When an arc is lost, power input is reduced until arcing is resumed. High electrode speed restores the arc quickly, resulting in higher average power input. Furnaces with high-speed electrodes make their heats more quickly than do low-electrode-speed furnaces in the same shops, scrap quality and other factors being equal.

T. R. Macon and **A. F. Sedgwick**: We appreciate the interest shown by the discussers and their constructive comments.

Mr. Concordia and Mr. Levoy are correct in their statement that the regulator will be significant in affecting voltage flicker only for those fluctuations that either occur slowly or are sustained for a few seconds.

Mr. Lloyd, in his discussion, mentioned that operation of the furnaces without the condenser would produce fluctuating voltages on the "borderline" or approaching "very objectionable" flicker. The power company felt that it was dangerous needlessly to create complaints about flicker and therefore have been reluctant to try operating without the condenser.

A Quick Method for Estimating Closed-Loop Poles of Control Systems

KAN CHEN
ASSOCIATE MEMBER AIEE

Synopsis: Present methods for evaluating closed-loop poles of feedback-control systems are often cumbersome and laborious when the order of the system increases beyond the third. For the purpose of saving the control engineers' time to a great extent, this paper presents a very quick method for estimating the closed-loop poles of feedback-control systems of any order. The results given by the method are approximate but have acceptable accuracy for all practical purposes. The method, as demonstrated in an example, forms the basis for a new system-synthesis technique.

ANALYSES and syntheses of feedback-control systems are carried out either in the frequency domain[1] or in the s-plane[2,3] where s is the complex variable in connection with Laplace transformation. Usually the s-plane approach leads to more direct and complete information about the system transient response. However, design in the frequency domain is relatively simple because the open-loop frequency characteristics can be quickly converted to the closed-loop frequency characteristics through the use of the "M-contours" or the "Nichol's chart,"[1] whereas no such straightforward technique exists for converting open-loop poles to closed-loop poles in the s-plane. The root locus method advocated by Evans[2] will place the closed-loop poles in evidence after the root loci are obtained. However, the construction of the root loci for a high-order system and the calibration of gain along the loci even for a low-order system are laborious. This paper presents a method which alleviates this difficulty of determining closed-loop poles from the given open-loop zeros and poles, thereby enhancing the value of the s-plane approach.

Accuracy Versus Simplicity

To fix ideas, consider the closed-loop system in Fig. 1, in which the open loop transfer function $G(s)$ is given in factored form. In other words, the zeros and poles of $G(s)$ are known. There is no problem in determining the zeros of the closed-loop transfer function $C(s)/R(s)$, since they are identical to the zeros of $G(s)$. The exact evaluation of the closed-loop poles of $C(s)/R(s)$, however, involves the solution of an nth order algebraic equation in s known as the characteristic equation of the closed-loop system. When the order of the system is above the third, time-consuming procedure of trial-and-error is usually used to solve for the exact closed-loop poles.

A number of approximate methods[4-6] are available for solving high-order algebraic equations. These methods, using a procedure of successive approximations, can yield results to any desired degree of accuracy. However, the first approximation yielded by these methods is generally too crude to be of any value. Excessive amount of labor is involved if accurate solution is desired. The problem is therefore one of accuracy of solution versus simplicity of method.

A numerical-reiteration process has been developed by G. A. Biernson[7,8] for computing the closed-loop poles based on the frequency asymptote plot of the open-loop transfer function G. The first or second approximation in this process yields very good results for the closed-loop poles corresponding to break frequencies distant from the crossover frequency, the frequency at which the asymptote of G crosses the 0-db (decibel) line. However, the problem of accuracy versus simplicity comes back when Biernson's method is used to extract the remaining closed-loop poles from the frequency asymptotes near the crossover frequency. Since the location of the

---

Paper 57-182, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted July 6, 1956; made available for printing January 3, 1957.

KAN CHEN is with Westinghouse Electric Corporation, East Pittsburgh, Pa.

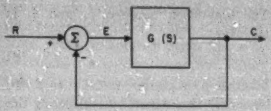


Fig. 1. Typical configuration of a feedback-control system

dominant closed-loop poles are closely related to the behavior of G near the crossover frequency, a quick and simple method for estimating these dominant closed-loop poles is particularly desirable.

The method presented by this paper extends Biernson's work[7,8] and allows one to arrive quickly at an approximate solution of *all* closed-loop poles with one try. Generally speaking, the purpose of evaluating the closed-loop poles is to give the control engineer a general idea of the system transient response during the process of design. Accurate response will be obtained by performing tests on the actual system or its computer representation. Extreme accuracy is therefore unimportant in evaluating the closed-loop poles. Furthermore, the parameters of a practical control system are seldom accurately known. Estimation of the closed-loop poles certainly need not be more accurate than the measurement of the parameters. That the proposed method yields acceptable accuracy for all practical purposes will be demonstrated in several examples.

Procedure of the Method

The proposed method is based upon a comprehensive use of the root locus plot[3] and the frequency asymptote plot.[1] As is well known to control engineers, the closed-loop poles will move in the s-plane from the open-loop poles to the open-loop zeros along the root loci, as the loop gain increases from zero to infinity. The gain information offered by the frequency asymptote plot enables one to determine quickly which closed-loop poles are near the beginning, which are near the end, and which are in the middle of migration along the root loci. By so dividing the closed-loop poles into three groups, their locations in the s-plane can be estimated with very little effort.

Consider that the open-loop transfer function $G(s)$ in Fig. 1 assumes the following form

$$G(s)=\frac{K(s+a_1)}{s(s+b_1)(s+b_2)(s+b_3)} \tag{1}$$

where $b_1<a_1<b_2<b_3$. The first step of the method is to plot the gain-frequency asymptotes as in Fig. 2. This can be done very quickly in all cases. In fact, if this system has been designed in the frequency domain, this plot is immediately available. The root loci of the closed-loop system are shown in Fig. 3(A) or (B), depending on the relative locations of the open-loop poles and zeros. They are shown here as a reference for later discussion, not as a necessary step of the proposed method.

Closed-Loop Poles Near the Beginning of Migration

For $|G|<<1$, the following approximation may be written,

$$\frac{C(s)}{R(s)}=\frac{G(s)}{1+G(s)}\approx G(s) \tag{2}$$

Under this condition, it is clear that the closed-loop poles of $C(s)/R(s)$ are approximately the open-loop poles of $G(s)$. Of course, it is difficult to draw a demarcation line for $|G|<<1$. This condition will be defined arbitrarily by $|G\ |\leqslant -15$ db, or $|G|\leqslant 0.178$. The choice of the 15-db line as the demarcation line is based on the circumspection that this choice will satisfy the condition $|G|\ll 1$ and will simplify the evaluation of closed-loop poles in the middle of migration along the root loci as will be explained in a later section.

Referring to Fig. 2, one approximate closed-loop pole is at $s=-b_3$ since this break on the frequency plot corresponds to a pole and is under the $|G|=-15$-db line. In terms of the root loci, this simply means that the closed-loop pole which starts from $-b_3$ has not moved far from $s=-b_3$ for the gain being considered. Since this closed-loop pole near $-b_3$ corresponds to a transient term that decays rapidly, it will not affect the over-all system transient very much if the closed-loop pole is considered as not having moved from $-b_3$ at all.

Closed-Loop Poles Near the End of Migration

For $|G|>>1$, the following approximation may be written,

$$\frac{E(s)}{R(s)}=\frac{1}{1+G(s)}\approx\frac{1}{G(s)} \tag{3}$$

Under this condition, it is clear that the closed-loop poles of $C(s)/R(s)$, which are the same as the poles of $E(s)/R(s)$, are approximately the open-loop zeros of $G(s)$. Again, for lack of a natural demarcation line for $|G|>>1$, this condition will be defined arbitrarily by $|G|\geqslant +15$ db, or $G|\geqslant 5.63$.

Referring to Fig. 2, one approximate closed-loop pole is near $s=-a_1$ since this break corresponds to a zero and is above the $|G|=+15$-db line. In terms of the root loci, this simply means that one closed-loop pole has moved very close to $-a_1$. Actually, however, this closed-loop pole has not quite reached $-a_1$. In fact, since $s=-a_1$ is a closed-loop zero, there is a "dipole" (formed by a closed-loop zero and a closed-loop pole close to each other) in the neighborhood of $s=-a_1$. The dipole has a significant effect on the settling time of the system since it corresponds to a small but long transient term and hence gives a "tail" to the system response.

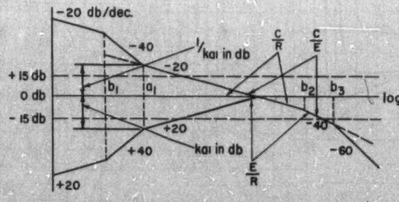
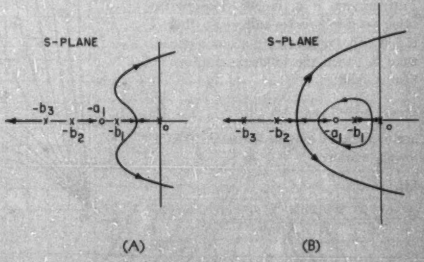


Fig. 2. The gain-frequency asymptote plot for equation 1



(A)

(B)

Fig. 3. The root loci corresponding to equation 1

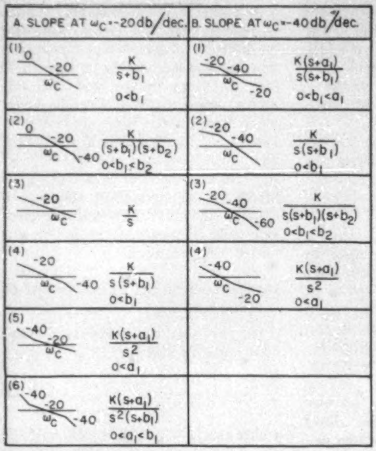

A. SLOPE AT ωc=-20db/dec.
B. SLOPE AT ωc=-40db/dec.
(1) 0 -20 ωc K/(s+b1) 0<b1
(1) -20 -40 ωc -20 K(s+a1)/s(s+b1) 0<b1<a1
(2) 0 -20 ωc -40 K/(s+b1)(s+b2) 0<b1<b2
(2) -20 -40 ωc K/s(s+b1) 0<b1
(3) -20 ωc K/s
(3) -20 -40 ωc -60 K/s(s+b1)(s+b2) 0<b1<b2
(4) -20 ωc -40 K/s(s+b1) 0<b1
(4) -40 ωc -20 K(s+a1)/s² 0<a1
(5) -40 -20 ωc K(s+a1)/s² 0<a1
(6) -40 -20 ωc -40 K(s+a1)/s²(s+b1) 0<a1<b1

Fig. 4. Common variations of the open-loop transfer function G near the crossover frequency ω_c

As a result of the foregoing discussion, the dipole should not be neglected by assuming the open-loop zero to be cancelled by the closed-loop pole for $|G|>>1$. Nevertheless, it is possible to determine the separation between the zero and the pole of the dipole quickly in an approximate manner. The residue of $E(s)$ at the pole of interest will be the magnitude of the corresponding transient term of both $c(t)$ and $e(t)$. The magnitude of this residue for a step input of position may be determined approximately as follows:

$$|k_{a1}|=\left|\frac{(s+a_1)}{s[G(s)+1]}\right|_{s=-a_1}\approx$$

$$\left|\frac{1}{sG(s)}\cdot(s+a_1)\right|_{s=-a_1}$$

$$=\left|\frac{(-a_1)(-a_1+b_1)(-a_1+b_2)(-a_1+b_3)}{K(-a_1)}\right| \tag{4}$$

Furthermore, if a_1 is not close to the other open-loop poles and zeros; that is, if a_1 is either much larger or much smaller than b_i; then the following approximation is valid:

$$|-a_1+b_i|\approx\sqrt{a_1^2+b_i^2} \tag{5}$$

and equation 4 may be written in the following manner:

$$|k_{a1}|\approx$$

$$\frac{\sqrt{a_1^2+0^2}\sqrt{a_1^2+b_1^2}\sqrt{a_1^2+b_2^2}\sqrt{a_1^2+b_3^2}}{K|a_1|} \tag{6}$$

The last expression is the approximate gain of $E(j\omega)/R(j\omega)$ at $\omega=a_1$ multiplied by $\sqrt{2}$, or simply the gain of the frequency asymptotes for $E(j\omega)/R(j\omega)$ at $\omega=a_1$. This value can therefore be read readily from the frequency asymptote plot. It is marked k_{a1} in Fig. 2.

It can be shown[3] that if the closed-loop pole $(-p_i)$ and the closed-loop zero $(-z_i)$ of a dipole are much closer to the origin than any other closed-loop poles and zeros, which is normally the case, the magnitude of the residue k_{pi} is given approximately by

$$k_{pi}=-1+\frac{p_i}{z_i} \tag{7}$$

The separation between p_i and z_i is therefore

$$p_i-z_i=k_{pi}z_i \tag{8}$$

Whether p_i is greater or less than z_i can be readily determined by considering the root loci which lie on the real axis. Thus, referring to Fig. 2, since $k_{pi}\approx k_{a1}$ and $z_i=a_1$, one closed-loop pole is quickly located at $s=-(1+k_{a1})a_1$ where k_{a1} is readily given by the frequency asymptote plot.

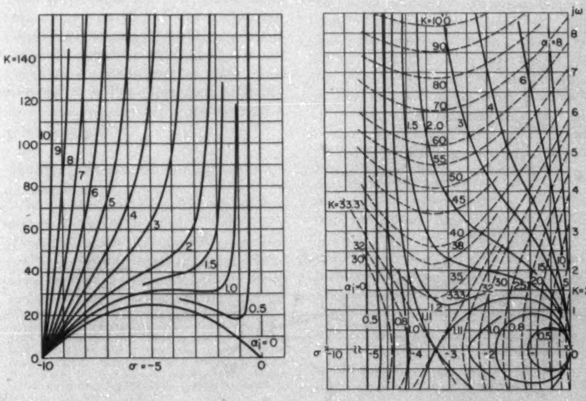
CLOSED-LOOP POLES IN THE MIDDLE OF MIGRATION

Considered so far have been the open-loop poles and zeros whose corresponding breaks on the frequency asymptote plot are outside the band of -15 db $<|G<$ $+15$ db. If attention is now focused on the breaks inside the ±15-db band and those outside this band ignored, the closed-loop poles in the middle of their migration may be approximately determined. As the number of breaks inside the ±15-db band is usually small, the problem of solving an nth order characteristic equation is greatly simplified. For instance, the open-loop transfer function $G(s)$ in Figs. 2 and 3 is now approximated by

$$G(s)=\frac{K(s+a_1)}{s(s+b_1)(s+b_2)(s+b_3)}\approx\frac{K/b_3}{s(s+b_2)} \tag{9}$$

since, for $|G|$ within the ±15-db band, it may be considered that $b_1<a_1<<|s|<<b_3$. The corresponding closed-loop transfer function is

$$\frac{G(s)}{1+G(s)}\approx\frac{K/b_3}{K/b_3+s(s+b_2)} \tag{10}$$



(A)

(B)

Fig. 5. Charts for graphical determination of closed-loop poles

The original quadratic characteristic equation has thus been reduced to a quadratic equation which may be easily solved. Note that the approximation given by equation 9 is equivalent to approximating the frequency plot in Fig. 2 outside of the ±15-db band by the slant dashed lines which are extensions of the actual asymptotes crossing the ±15-db lines.

Of course, approximation within the ± 15-db band does not result in a quadratic characteristic equation for every system. However, if the system has been properly designed for stability, complicated as the system may be, the frequency asymptotes near the crossover frequency ω_c usually assume one of the forms given in Fig. 4. As shown in Fig. 4, the resultant characteristic equation will be at worst a cubic. Furthermore, these cubic open-loop transfer functions will assume either the form

$$\frac{K(s+a_1)}{s^2(s+b_1)} \text{ or the form } \frac{K}{s(s+b_1)(s+b_2)}$$

To facilitate solving for their corresponding closed-loop poles, Fig. 5(A) and 5(B) (taken from reference 9) give the real and complex poles, respectively, corresponding to

$$G(s)=\frac{K(\lambda+\alpha_t)}{\lambda^2(\lambda+10)}$$

where K and α_t are parameters. The s-plane has been normalized and has become the λ-plane where $\lambda=\sigma+j\omega$, so that the simple pole is at $\lambda=-10$. Fig. 6(A) gives the real poles and Figs. 6(B) and (C) (taken from reference 9) give the complex poles corresponding to the following equation

$$G(s)=\frac{K}{(s+\sigma_{12}-j\omega_{12})(s+\sigma_{12}+j\omega_{12})(s+\sigma_3)}$$

where K and $(\sigma_3-\sigma_m)$ are parameters, ω_{12} may be real or imaginary, and σ_m is defined in Fig 6(A). Again the s-plane has been normalized so that the first two poles are at either $(\sigma-\sigma_m)=\pm 1$ or $\omega_{12}=\pm 1$. The use of some of these figures will be illustrated later in the examples.

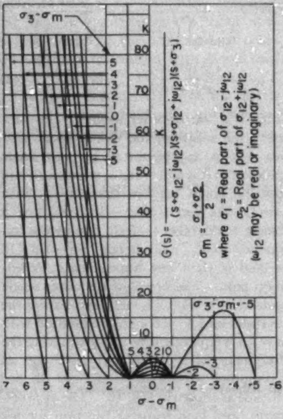
In some cases, the open-loop frequency asymptotes within the ±15-db band may not be described by any of the variations in Fig. 4. The frequency asymptote shown in Fig. 7 represents such a case. A double-order open-loop pole at $s=-b_1$ changes the slope of the frequency asymptotes from -20 db per decade to -60 db per decade. However, the frequency asymptotes within the ±15-db band may be reduced to one of the forms given in Fig. 4 if the double-order pole at $-b_1$ is approximated by two single poles at $-\alpha b_1$ and $-b_1/\alpha$ so that now the break at $\omega=\alpha b_1$ is outside the ±15-db band. Chestnut and Mayer have shown[10] that this approximation will not affect the closed-loop system behavior substantially if the factor α is not much greater than two. This kind of approximation makes the proposed method applicable to practically all closed-loop systems which have been properly designed for stability without having to solve a cubic or higher order equation. At any rate, the idea of the proposed method is to reduce the number of essential closed-loop poles in the middle of migration to such an extent that their approximate location can be evaluated by the simplest root-locus construction.

It should be remarked that the proposed method is also useful for factoring a polynomial of the form

$$K\prod^{i}(s+a_i)+\prod^{j}(s+a_j)$$

which may arise from time to time in simplifying the transfer functions of multiloop servo systems. The factorization problem is simply the same as finding the closed-loop poles for a servo system with the open-loop transfer function

$$G(s)=K\prod^{i}(s+a_i)/\prod^{j}(s+a_j)$$



(A)

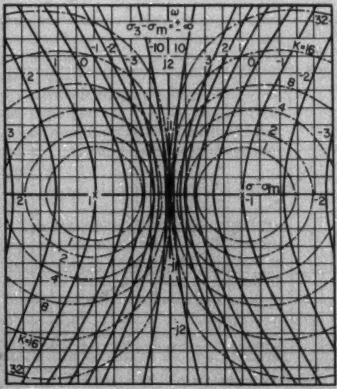
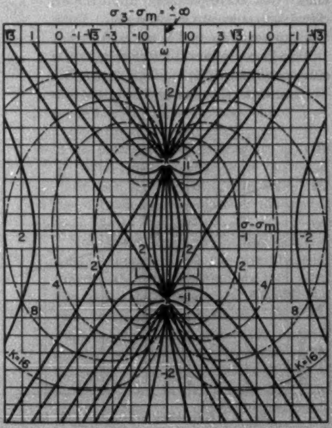


Fig. 6 (above, left, and right). Charts for graphical determination of closed-loop poles

(B)



(C)

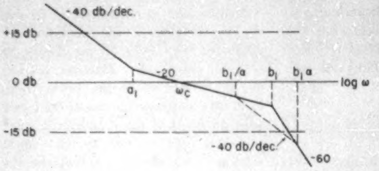


Fig. 7 (left). An unusual case of the open-loop frequency asymptotes near the crossover frequency

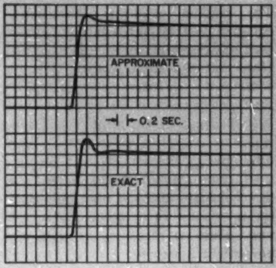


Fig. 9 (right). Approximate and exact transient responses to a step input for example 1

Examples

Example 1

The closed-loop poles of a system with the following open-loop transfer function are to be determined.

$$G(s)=\frac{20(s+1)}{s\left(\frac{1}{0.5}s+1\right)\left(\frac{1}{20}s+1\right)\left(\frac{1}{50}s+1\right)}\tag{11}$$

The corresponding frequency asymptotes are first plotted as in Fig. 8. The break at $\omega=50$ is under the -15-db line, so one approximate closed-loop pole is at $s=-50$. The break at $\omega=1$ is above the $+15$-db line and the gain at this point is 20 db or 10, so another approximate closed-loop pole, which is to the left of -1 in the s-plane, is at $s=-1(1+1/10)=-1.1$. For $|G(\omega)|$ within the ±15-db band,

$$G(s)\approx\frac{20(s)}{s\left(\frac{1}{0.5}s\right)\left(\frac{1}{20}s+1\right)(1)}=\frac{10}{s\left(\frac{1}{20}s+1\right)}\tag{12}$$

and

$$\frac{G(s)}{1+G(s)}\approx\frac{10}{s(0.05s+1)+10}\tag{13}$$

Setting the denominator to zero, one finds the corresponding approximate closed-loop poles at $s=-10\pm j10$. These four approximate poles at -50, -1.1, $-10\pm j10$ compare rather closely to the exact poles at -55.2, -1.06, $-7.15\pm j11$. The transient responses of the system to a step, based on these two sets of closed-loop poles, are obtained from an analog computer and are shown in Fig. 9. As seen, the delay time, the rise time, and the time at which maximum overshoot occurs are almost identical on the two curves. The overshoot is 12% for one and 17% for the other. Both transients have about the same long tail. It is felt that these two curves are close enough to each other for practical transient evaluation, yet the amount of work in finding the closed-loop poles has been reduced a great deal by using the proposed method.

Example 2

The closed-loop poles of a system with the following open-loop transfer function are to be determined. (This is an example given in Biernson's paper.[8])

$$G(s)=\frac{500\left(\frac{s}{0.2}+1\right)(s+1)}{s\left(\frac{s}{0.04}+1\right)^2\left(\frac{s}{16}+1\right)^2}\tag{14}$$

The corresponding frequency asymptotes are first plotted in Fig. 10. The break at $\omega=0.2$ corresponding to an open-loop zero is above the $+15$-db line and the gain at this point is 40-db or 100. So an approximate closed loop pole which is to the right of -0.2 in the s-plane is at $s=-0.2(1-1/100)=0.198$. For $|G|$ within the ±15-db band, the open-loop transfer function is approximately given:

$$G(s)\approx\frac{500\left(\frac{s}{0.2}\right)(s+1)}{s\left(\frac{s}{0.04}\right)^2\left(\frac{s}{16}+1\right)^2}=\frac{4(s+1)}{s^2\left(\frac{s}{16}+1\right)^2}$$

$$=\frac{4\times16^2(s+1)}{s^2(s+16)^2}\tag{15}$$

This is a fourth-order system. To simplify the solution, the double-order pole at -16 is approximated by two single poles at -16α and $-16/\alpha$, where α should not be much greater than two. With $\alpha=1.6$ the single-pole -16α will give a break under the -15-db line in the frequency plot, and hence $s=-1.6\times16=25.6$ is then one approximate closed-loop pole. The break of the other single pole at $s=-16/1.6=-10$ is still inside the ±15-bd band. This approximation with $\alpha=1.6$ is equivalent to approximating the frequency asymptotes near $\omega=16$ by an asymptote with a slope of -40 db per decade as shown in Fig. 10. The approximate open-loop transfer function for $|G|$ within the ±15-db band is now given by

$$G(s)\approx\frac{4\times16^2(s+1)}{s^2(s+10)25.6}=\frac{40(s+1)}{s^2(s+10)}\tag{16}$$

The chart of Fig. 5(B) may then be used

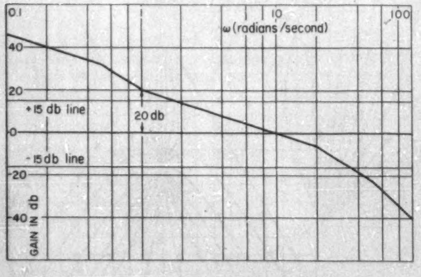


Fig. 8. Frequency asymptote plot for example 1

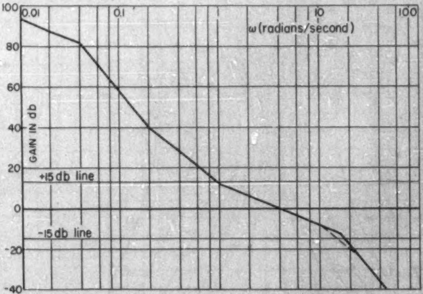


Fig. 10. Frequency asymptote plot for example 2

Vol.
Pt. 2
1957

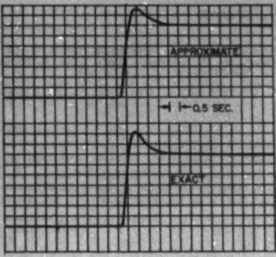


Fig. 11 (left). Approximate and exact transient responses to a step input for example 2

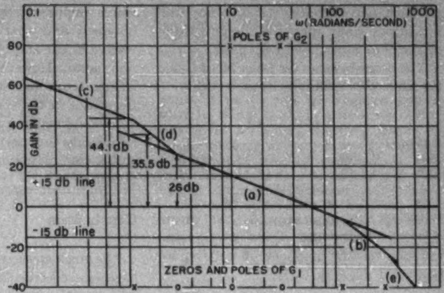


Fig. 13 (right). Construction of the frequency asymptotes for the design problem in example 3

to determine the corresponding closed-loop poles. Since the single pole in equation 16 is at $s=-10$, Figs. 5(A) and (B) may be used directly without normalization. The three closed-loop poles are read from the charts, with $K=40$ and $\alpha_i=1$. They are: $s=-1.45$ from Fig. 5 (A) and $s=-4.35\pm j3.15$ from Fig. 5(B).

The five closed-loop poles obtained by the proposed method are at

$$s=-0.198;\ -1.45;\ -4.35\pm j3.15;\ -25.6$$

These may be compared with the exact poles at

$$s=-0.198;\ -1.51;\ -4.35\pm j3.3;\ -22.6^{*}$$

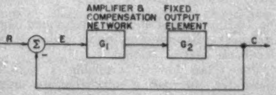


Fig. 12. Configuration of the feedback-control system in example 3

The corresponding system transient responses to a step input are shown in Fig. 11. As seen, the result given by the proposed method is again acceptable for practical purposes. If a value other than 1.6 had been chosen for α, the approximate poles at $s=-25.6$ and $-4.35\pm j3.15$ would assume different values. However, the corresponding system transient response would still be very much like that in Fig. 11 as long as α is not much greater than two.

EXAMPLE 3

This example is given to show that the proposed method and accessory charts may be used in design. The approach of design to be used is that promoted by Truxal[3]; namely, to force the system to meet the specifications in a straightforward synthesis procedure. Consider the following specifications for the system in Fig. 12.

1. The fixed part of the system is described by the transfer function

$$G_2(s)=\frac{24}{s\left(\frac{s}{10}+1\right)\left(\frac{s}{30}+1\right)} \qquad (17)$$

2. The performance characteristics are specified by the bandwidth, velocity constant, and relative damping ratio associated with the dominant closed-loop poles:

Bandwidth: about 60 radians per second
Velocity constant: about 160 per second
Damping ratio: approximately 0.7 for the dominant poles

The synthesis procedure starts from choosing one of the forms given in Fig. 4 for the frequency asymptotes near the crossover frequency, which is 60 radians per second. Suppose one chooses form 4 in Fig. 4(A). Then a −20 db per decade asymptote (*a*) will be drawn through the crossover point; viz, 0 db at 60 radians per second, as in Fig. 13. The value of b_1, at which the break within the ±15-db band occurs, is to be chosen. Extending the asymptote (*a*) to $\omega=1$ radians per second, one gets the gain of 35.5 db or 60 for the approximate $G_1G_2(s)$ within the ±15-db band. (The gain of 60 is also given directly by the crossover frequency of 60 radians per second since the asymptote (*a*) represents the transfer function $K/j\omega$.) The desirable form of $G_1G_2(s)$ within the ±15-db band is therefore

$$G_1G_2(s)\approx\frac{60}{s\left(\frac{s}{b_1}+1\right)} \qquad (18)$$

or

$$\frac{G_1G_2}{1+G_1G_2}\approx\frac{60}{\frac{s^2}{b_1}+s+60} \qquad (19)$$

and the damping ratio of the dominant closed-loop poles is given by

$$\zeta=\frac{1}{2\sqrt{\frac{60}{b_1}}} \qquad (20)$$

Setting ζ equal to the specified value of 0.7, the value of b_1 is determined to be 118.

Now the system described by the two frequency asymptotes (*a*) and (*b*) with a break at $\omega=118$ will have a velocity constant of 60 radians per second. In order to raise the velocity constant to 160 radians per second as specified, the −20 db per decade asymptote in the low-frequency region should go through the point 44.1 db at 1 radian per second. The low-frequency asymptote (*c*) is accordingly drawn as shown in Fig. 13. Asymptotes (*a*) and (*c*) can be linked by another asymptote (*d*) with a slope of −40 db per decade. It is clear that the asymptote (*d*) can be anywhere as long as its associated break points do not occur within the ±15-db band so as not to affect the damping ratio of the dominant poles. However, in order that the slow transient term corresponding to the closed-loop pole near the open-loop zero introduced by the asymptote (*d*) be not greater than 5% or 1/20 of the step input; the frequency plot at the break between asymptotes (*a*) and (*d*) should have a gain value of 20, or 26 db (see equation 6). This consideration fixes the location of the asymptote (*d*) as shown in Fig. 13.

Now if the desirable transfer function G_1G_2 is compared with the given transfer function G_2, it is clear that the two open-loop poles of G_2 have to be cancelled by two zeros in G_1. In order that the function of G_1 can be physically realized, another pole has to be included in G_1. Thus the asymptote of G_1G_2 is allowed to have another break at $\omega=300$. The reason a break is chosen at this frequency is that the break is now below the ±15-

---

* These exact poles were evaluated by Biernson.[7] In the same paper Biernson proposed an approximate method by which he arrived at almost the exact real poles after three or four iterations, and used a graphical procedure to estimate the complex poles to be $s=-4.03\pm j3.88$.

db band and therefore its existence will not substantially affect the damping ratio of the dominant closed-loop poles.

The desirable transfer function of $G_1G_2(s)$ can now be readily written:

$$G_1G_2(s)=\frac{160\left(\frac{s}{3}+1\right)}{s\left(\frac{s}{1.12}+1\right)\left(\frac{s}{118}+1\right)\left(\frac{s}{300}+1\right)} \tag{21}$$

Dividing this expression by the given $G_2(s)$, provides the desirable transfer function for the amplifier and the compensation network $G_1(s)$:

$$G_1(s)=\frac{6.67\left(\frac{s}{3}+1\right)\left(\frac{s}{10}+1\right)\left(\frac{s}{30}+1\right)}{\left(\frac{s}{1.12}+1\right)\left(\frac{s}{118}+1\right)\left(\frac{s}{300}+1\right)} \tag{22}$$

The compensation network can then be synthesized by resistance-capacitance components through the modern network synthesis technique.[11]

The proposed synthesis procedure utilizing the quick method of estimating the closed-loop poles had the advantages that the desirable transient response, which implies stability, can be directly obtained and that the effort normally wasted in cut-and-try is reduced to nil or minimum. However, the brute-force procedure tends to overlook inherent nonlinearities in the control system and the system arrived at by the procedure is not necessarily the most economical from the standpoint of number and size of components.

Conclusions

The quick method for estimating closed-loop poles described in this paper emerged from an attempt to combine the advantages of the frequency-response approach and the root-locus approach to the analysis of feedback-control systems. The approximate solutions provided by the method have acceptable accuracy for the practical purpose of evaluating the essential characteristics of the transient as well as the steady-state responses of a closed-loop system. As a result of its simplicity, the method provides the basis for a new direct synthesis procedure for linear control systems as was demonstrated in the paper.

References

1. Principles of Servomechanisms, G. S. Brown, D. P. Campbell. John Wiley & Sons, Inc., New York, N. Y., chap. 8, 1948.

2. Control-System Dynamics, W. R. Evans. McGraw-Hill Book Company, Inc., New York, N. Y., 1954.

3. Automatic Feedback Control System Synthesis, J. G. Truxal. McGraw-Hill Book Company, Inc., chaps. 5 and 6, 1955.

4. Mathematics of Modern Engineering, R. E. Doherty, E. G. Keller. Vol. 1, John Wiley & Sons, Inc., 1936, pp. 98–128.

5. A Method of Successive Approximations of Evaluating the Real and Complex Roots of Cubic and Higher-Order Equation, S. N. Lin. *Journal of Mathematics and Physics*, Cambridge, Mass., vol. 20, Aug. 1941, pp. 231–42.

6. A Method of Finding Roots of Algebraic Equations, S. N. Lin. *Ibid.*, vol. 22, June 1943, pp. 60–77.

7. Quick Methods for Evaluating the Closed-Loop Poles of Feedback Control Systems, George Biernson. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 53–70.

8. A General Technique for Approximating Transient Response from Frequency Response Asymptotes, George A. Biernson. *Ibid.*, vol. 75, pt. II, Nov. 1956, pp. 253–73.

9. Synthesis of Feedback Control Systems by Gain-Contour and Root-Contour Methods, Victor C. M. Yeh. *Ibid.*, May 1956, pp. 85–96.

10. Servomechanisms and Regulating Systems, H. Chestnut, R. W. Mayer. Vol. 1, John Wiley & Sons, Inc., 1951, pp. 415–16.

11. Synthesis of RC Networks, H. J. Smead. *Engineering Bulletin*, Purdue University, Lafayette, Ind., vol. 39, no. 1, Jan. 1955.

◆

Discussion

G. W. Bills (North American Aviation, Inc., Downey, Calif.): The ideas presented in this paper are interesting and practical, and should help systems engineers in the feedback-control field. While it is true that ordinarily the solution of the characteristic closed-loop equation for exact closed-loop poles whether by root-locus construction or trial-and-error procedures such as Lin's (see references 5 and 6 of the paper) are time-consuming, with the aid of the general-purpose digital computer such solutions are quick, accurate, and routine.

At North American Aviation we are programming root-locus solutions on the International Business Machines Corporation (IBM) *704* and *701*. This is being done by root-extraction routines and the answers will be produced in the form of a root locus already plotted and photographed complete with scales and titles.

Transient responses will also be computed by the IBM's employing the Boxer and Thaler technique or a Runge-Kutta method for solving differential equations. The results of these calculations will also be presented in the form of curves as well as numerical data.

It is believed that the two sets of curves will be adequate for synthesis procedures. Although at present we have not had sufficient experience to evaluate the cost per set of curves as produced by an IBM versus that produced by an engineer and a slide rule or desk calculator, we believe that cost figures will favor the digital computer. An additional advantage of the computers is that it will require less engineering brain power of which there is a shortage.

However, in instances where general-purpose digital computers are not available or where only a few closed-system analyses are required and to write a general program would be too costly, the method of this paper offers definite advantages as compared to older methods.

Casimir J. Doda (North American Aviation, Inc., Downey, Calif.): That the knowledge of the location of the closed-loop poles is of primary importance in determining the performance characteristics of a feedback-control system will not be disputed. The author's method for quickly estimating the closed-loop poles of control systems, therefore, is a valuable contribution to the field and he should be congratulated for its presentation.

Much more can be said in favor of sketching in with rapidity and yet with some degree of accuracy the root-locus plot of the open-loop transfer function. This can be done by starting with the asymptotic relationships and then further considering the separation of the poles and zeros and the effects of dipoles to sketch in the various curvatures. Considerations of limiting conditions as two sets of widely separated poles and zeros are further separated will quickly locate the limits of excursion of the loci.

Aided by such a sketch the approximations obtained by Mr. Chen's method can be further refined before determining the transient response and performance of a feedback control system.

For those having available a digital computer, the road of approximation may not be a road of as great enchantment as actually determining the root-locus plot digitally, having at the same time gain values on the locus plot. At the present time we are optimistic about programs to accomplish this. We had hoped to have results by now but this needs to wait for additional checkout.

George Biernson (Sylvania Electric Products, Inc., Waltham, Mass.): Mr. Chen has presented a noteworthy simplification to the problem of calculating the closed-loop poles of feedback-control systems. The approximation technique I originally presented is quite effective for calculating poles in the frequency regions where the loop gain is either much greater than unity or much less than unity, but is rather cumbersome in the region of unity gain. Mr. Chen has found a simple technique for calculating the poles in that region of unity gain.

To supplement the simplification that Mr. Chen has presented for calculating closed-loop poles in the region of unity loop gain, I should like to describe some approximations for further simplifying the calculation of closed-loop poles in the regions of high-loop gain and of low-loop gain. These approximations are presented in more detail in reference 8 of the paper.

In essence, the approximation rules are as follows:

1. In the frequency region where the loop gain is much greater than unity, the closed-loop poles are approximately equal to the open-loop zeros.
2. In the frequency region where the loop gain is much less than unity, the closed-loop poles are approximately equal to the open-loop poles.
3. A closed-loop pole is shifted from the corresponding open-loop zero or pole into the frequency region where the asymptote slope of the open-loop transfer function is an odd power of frequency (i.e., has a slope which is an odd multiple of −20 db per decade).
4. Where the loop gain is much greater than unity, the magnitude of the complex-frequency shift of the closed-loop pole from the open-loop zero is approximately equal to the magnitude of that zero divided by the asymptote of the loop gain at that zero.
5. Where the loop gain is much less than unity, the magnitude of the complex-frequency shift of the closed-loop pole from the open-loop pole is approximately equal to the magnitude of the open-loop pole multiplied by the asymptote of the loop gain at the open-loop pole.

To illustrate these rules, assume that the open-loop transfer function consists only of real poles and zeros:

$$G(s)=\frac{\pi(s+\alpha)}{\pi(s+\beta)}$$

Assume that a zero $-\alpha_1$ occurs where the loop gain $|G(j\omega)|$ is much greater than unity. Since there must be a closed-loop pole nearby, designate this pole as s_{c1}. Designate δ_1 as the complex-frequency shift of s_{c1} from $-\alpha_1$, i.e.:

$$s_{c1}=-\alpha_1+\delta_1$$

By rule 4 the magnitude of δ_1 is approximately

$$|\delta_1|\approx|-\alpha_1|/asm|G(-\alpha_1)|$$

where $|-\alpha_1|$ is the magnitude of the open-loop zero, and $asm|G(-\alpha_1)|$ represents the magnitude asymptote of G at the frequency $|-\alpha_1|$. If the asymptote slope of G above the frequency of the open-loop zero is an odd multiple of −20 db per decade, the closed-loop pole s_{c1} is shifted toward increasing frequency, and it is approximately equal to

$$s_{c1}=-\alpha_1-|\delta_1|\approx-\alpha_1\left[1+\frac{1}{asm|G(-\alpha_1)|}\right]$$

If the asymptote of G below the open-loop zero is an odd multiple of −20 db per decade, the closed-loop pole s_{c1} is shifted toward decreasing frequency, and it is approximately equal to

$$s_{c1}=-\alpha_1+|\delta_1|\approx-\alpha_1\left[1-\frac{1}{asm|G(-\alpha_1)|}\right]$$

Similar expressions can be derived for the closed-loop poles in the region where the loop gain is much less than unity. The magnitude of the shift δ_2 of a closed-loop pole s_{c2} from a corresponding open-loop pole $-\beta_2$, occurring in the region where the loop gain is much less than unity, is approximately (by rule 5)

$$|\delta_2|\approx|-\beta_2|asm|G(-\beta_2)|$$

I should like to close with a note of caution concerning the use of the closed-loop poles in feedback-control design. To define the transient response, one must know both the closed-loop poles and the coefficients. In those systems (which are quite common) where there are more than two closed-loop poles in the frequency region of unity open-loop gain, specifying the damping ratio and natural frequency of the so-called "dominant pair" of poles does not set a very reliable limitation upon the transient performance.

I am a firm adherent to the use of the frequency response as the fundamental tool for feedback-control design. Probably the main reason why there is so much effort to use the actual poles in synthesis is that there is not adequate understanding as to how frequency-response characteristics can be interpreted in terms of transient performance. In fact, the error-versus-input frequency-response characteristic (which is much more valuable for estimating transient performance than the return-signal-versus-input frequency response) is hardly touched in the literature. The main value from studying the closed-loop poles is to provide insight to bridge the gap between frequency response and transient response. It may also be desirable to determine the closed-loop poles as a final check on design. However, the use of the closed-loop poles themselves as the primary frame of reference in design appears to add excessive complexity to the problem, and to obscure many of the important physical considerations involved. Besides, more reliable estimates of transient performance can be obtained from frequency-response data than from examining the values of the closed-loop poles and zeros.

Kan Chen: I wish to thank the discussers for their valuable comments. It is interesting to learn from Mr. Bills and Mr. Doda that digital computers are being investigated as a practical means for aiding the solution of the root loci. Without doubt this will be a great help to the control designer if the use of the digital computer for such purposes is justified economically. However, it should be pointed out with caution that the value of the computer as a design tool should not be overemphasized. A computer can do only what the engineer wants it to do and the engineer knows what he wants to do in a design problem only if he has gained a certain degree of insight to the problem. It is felt that the approximation method proposed in this paper has its chief value in providing this insight to the design engineer.

The five approximation rules in Mr. Biernson's discussion, especially the third one, are valuable supplements to the proposed method for estimating closed-loop poles near the beginning and near the end of migration along the root loci. I believe, however, that the fifth rule is seldom necessary for two reasons. First, the amount of shift of the poles near the beginning of migration is usually very small; and second, since these poles correspond to rapidly decaying transient terms, it will not affect the over-all system transient very much if these poles are considered as not having moved at all.

I agree with Mr. Biernson that in many practical systems there may be more than two dominant closed-loop poles. In this case, the effect of the closed-loop pole-zero configuration on the transient response may not be immediately apparent. It is recalled that the study on this effect, which has been done by Mulligan,[1] Chu,[2] and others was based on the assumption of only two dominant closed-loop poles. However, it should be pointed out that in synthesis the designer has the freedom to pick the number and the location of the dominant closed-loop poles. There is no reason why he cannot choose only a pair of dominant poles as was done in example 3 of the paper, and then the specification on the location of the dominant poles will have significant meanings. For this reason the proposed method is perhaps even more useful for synthesis than for analysis of closed-loop systems.

There are many pros and cons of the frequency response method versus the root locus method as a design approach. I wish to point out again that the proposed method is based on an attempt to combine the advantages and discard the disadvantages of the two afore-mentioned approaches. It is felt that more can be done along this direction.

REFERENCES

1. THE EFFECT OF POLE AND ZERO LOCATIONS ON THE TRANSIENT RESPONSE OF LINEAR DYNAMIC SYSTEMS, J. H. Mulligan, Jr. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 37, 1949, pp. 516–29.

2. SYNTHESIS OF FEEDBACK CONTROL SYSTEM BY PHASE-ANGLE LOCI, Yaohan Chu. *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 330–39.

Evaluation of Feedback-Control Systems Subjected to Large Signals

VINTON B. HAAS, JR.
ASSOCIATE MEMBER AIEE

IN THIS PAPER methods are presented for determining design modifications to improve the large-step response of a feedback-control system. This may be the response to a command or to a disturbance. The methods are applicable to systems of higher order as well as second order, and to systems with one or more important nonlinear effects. It is shown that the results obtained by application of the methods may be used to investigate questions not ordinarily answered by studies on a system subjected to small signal variations, such as:

1. Should a controller which has a dual mode of operation be instrumented for the system?
2. Should the saturation level for the cascaded components of the system be increased, thus usually increasing the linear range; or decreased, thus usually saving weight, bulk, and cost.
3. For an over-all required power gain in several cascaded components from the error to output, how should this power gain be distributed among the several cascaded components?

In feedback-control systems it is often found that a controller which has a dual mode of operation offers definite advantages.[1–3] (The three references given account only for the early work on this subject. There have been innumerable other contributions in this area.) Such a controller has one mode of operation for small signals and a second mode for large signals. The first mode can be designed from linear studies using a periodic, aperiodic, or statistical test signal, depending on the application. A large-step test function is usually satisfactory, regardless of the application, for design of the second mode.

A controller has two important properties which affect the large-step response of the system:

1. The maximum signal output level of the controller.
2. The power gain, signal gain, and associated response time of the controller.

In order to obtain design information relative to these properties and also to determine the improvement that may be obtained by instrumenting a second controller mode for a given system, the following three items regarding system performance are obtained by the methods of this paper:

1. The large-step response of the system when only a single-mode controller is used.
2. The large-step response of the system with an "optimum" dual-mode controller, based upon the neglecting of the controller's dynamic properties (assuming the controller has zero response time and infinite gain).
3. The large-step response of the system with an "optimum" dual-mode controller where the dynamic properties of the controller are considered.

The difference between responses 2 and 3 is due to the dynamics of the controller. The difference between responses 1 and 3 is due to the difference between a single-mode controller and a dual-mode controller.

The usefulness of these three items is enhanced by the fact that the methods given in this paper for obtaining these items require the performing of only three relatively simple calculations, described here as the three "evaluation tests." These three tests are first explained, and it is then shown how they may be used to provide the information listed. Finally, the interpretation of these items in terms of design is discussed.

The Three Evaluation Tests

The three evaluation tests to be described are all for a large step of the same amplitude. It is therefore important at the outset to determine what is to be considered as a "large step." Although no definite rule can be laid down in this regard, it is suggested that the amplitude of the step correspond to a possible situation that may be encountered during actual operation of the system, and that it should be sufficient to result in controller saturation. If difficulty is encountered in selecting a single large-step test function, it is appropriate to use several of different amplitudes and to apply the evaluation tests for each amplitude.

TEST 1: THE FORCING-ABILITY TEST

In a feedback-control system the controller has an algebraic maximum and minimum value of signal output that can be handled, as shown in the signal-flow diagram of Fig. 1. The forcing-ability test consists of assuming that it is possible for the controller output signal to vary between its maximum and minimum levels in a negligible time interval and then determining what the optimum transient response of the system would be assuming an "optimum" output signal from the controller. This optimum controller signal is that which will accomplish the specified desired change of system output in a minimum time without overshoot. The result of interest is the time duration of the optimum transient response of the system. This time duration is described here as the "forcing-ability contribution" of the system. The determination of this contribution, expressed in units of time, is relatively simple because only the open-loop characteristics of the system need be considered, as shown in Fig. 2. In the computation, all nonlinear characteristics of the output device must be considered. When the value of the forcing-ability contribution is thus obtained, it is a measure of the ability of the output device to be forced to accomplish desired large-step changes.

TEST 2: FORCING-UTILIZATION TEST

The forcing-utilization test is performed by introducing a fictitious external signal of constant magnitude to the point on the signal-flow path indicated by Fig. 3. First the appropriate value of this external signal must be established; this is done as follows: It is considered that the closed-loop system has first been subjected to the specified large-step input and has attained its steady-state response. A trial magnitude of external signal is then applied resulting in another corrective transient; after this latter transient dies away the output of the controller, ψ, will have been modified. The particular magnitude of ex-

---

Paper 57-201, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 22, 1956; made available for printing November 30, 1956.

VINTON B. HAAS, JR., is with the University of Connecticut, Storrs, Conn.

This paper is taken, in part, from a thesis submitted to the Electrical Engineering Department at the Massachusetts Institute of Technology. The writer wishes to acknowledge the guidance of Professors S. J. Mason, D. C. White, and A. Kusko who supervised the thesis. The work was conducted by the Energy Conversion Group of the Servomechanisms Laboratory at the Massachusetts Institute of Technology and was supported, in part, by Wright Air Development Center under Contract *AF33-(616)3243*. Professor R. J. Kochenburger, University of Connecticut, gave many helpful suggestions during the writing of the paper.

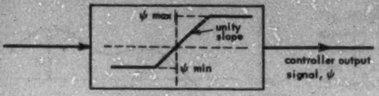


Fig. 1. Representation of the maximum and minimum controller output signal

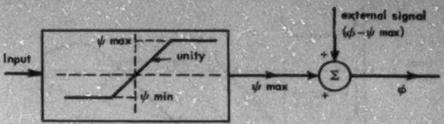


Fig. 3. Configuration for the forcing utilization test

ternal signal that results in this final controller output, ψ, being just equal to its maximum magnitude value, ψ_{max}, is the appropriate magnitude to be employed in the test.

In the event that either the forcing signal or the output device is not symmetrical, there are two parts to this test: (1) the utilization contribution for the algebraic maximum forcing is found, and (2) the utilization contribution for the algebraic minimum forcing is found. To find the latter, the external signal is adjusted to give the algebraic minimum controller output in the steady state.

The actual test consists of removing this external signal after it has previously been applied sufficiently long enough for steady-state response to have been attained. The time duration of this transient response is the result of interest and is referred to as the "forcing-utilization contribution." The time duration associated with this second test is defined in terms of a tolerance band about the final value. For example, a band of 2 per cent would mean that the transient is considered as completed when the transient error becomes and remains less than ±2 per cent. This same tolerance band used in defining the second contribution must also be used later to define the duration of the actual large-step response. This contribution is a measure of the ability of the system to utilize the forcing signal that is available from the controller. Ideally, the system should exhibit no output response for the forcing-utilization test; practically, only a response of small magnitude and short duration should be obtained. Thus, in most cases the system will operate essentially linearly during this transient, and therefore a linear approximation of any system nonlinearities affecting large signal behavior is justified for making the calculation. The controller itself will appear linear since all operation during this test is along the straight-line portion of its characteristic. Although the calculation is for a closed-loop transient, it is simplified by the fact that linear procedures may be used.

TEST 3: MAGNITUDE OF THE FIRST RESPONSE OVERSHOOT

This test consists of determining the magnitude of the first overshoot in response to the large-step input applied to the system with the single-mode controller. From an analytical point of view, the calculations of tests 1 and 2 are relatively much easier than in this third test. This is fortunate, since it is usually required that tests 1 and 2 be performed more times in an evaluation than test 3.

A special interpretation is placed on the definition of the magnitude of the first overshoot. x represents the error. The "first overshoot" is defined as the value of that overshoot which occurs when $\dot{x} = dx/dt$ first becomes zero after the step has been applied. Mathematically, negative overshoots are considered as zero overshoot. Hence the response shown in Fig. 4 represents a case of zero first overshoot while the response of Fig. 5 represents a finite first overshoot of value a. In most cases this first part of the transient is not too difficult to calculate even though the calculation is for a closed-loop nonlinear system.

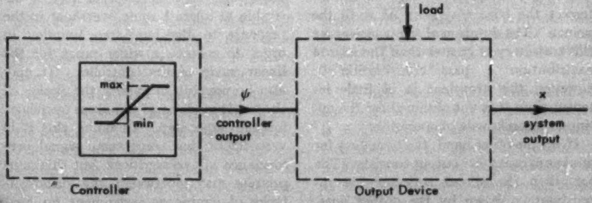


Fig. 2. The open-loop element for the forcing-ability test

Using the Results of the Three Evaluation Tests to Obtain the Three Desired Items

The first evaluation test, the forcing-ability test, is performed to obtain the system large-step response when the algebraic maximum and minimum controller-output signals are utilized in an optimum manner. The mechanics of utilizing the controller-output signal in an optimum manner are effectively demonstrated by the use of the phase plane. Fig. 6 gives an example of the large-step trajectory with a second-order output device when the controller-output signal is utilized in an optimum manner. The diagram is plotted so that the origin represents the desired condition for the system output, and so that owing either to a large-step command or to the application of a large-step disturbance, the system output is initially displaced from the origin. In order to correct the error in the system output, the maximum controller-output signal is initially applied to the output device. If something less than maximum is applied, the trajectory might be as shown by the dotted line, and the response time is longer. If the controller maximum output signal were to be maintained, a point would be reached at which x would be zero but not $\dot{x}$; thus the transient would not be completed. Therefore, at a certain optimum instant, the controller output is reversed from maximum output to minimum output so that, while the controller output is minimum, the trajectory passes through the origin, completing the transient. If reversal occurs before this optimum instant, the transient is unnecessarily long; if reversal occurs afterward, a transient overshoot results and the transient is, again, unnecessarily long. The trajectory shown in Fig. 6 is mathematically the required trajectory to obtain the minimum time of response due to the large step.

The trajectory corresponding to the minimum controller-output part of test 1 is plotted as a family of curves as represented by *qo* and *ba* in Fig. 7. Test 2 establishes a specific value of output error, x, at which, for $\dot{x}=0$, the control system

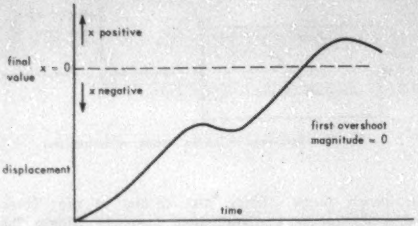


Fig. 4. **Special case for the first overshoot magnitude where first overshoot is zero**

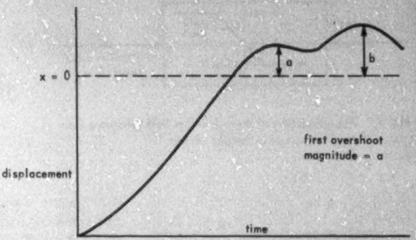


Fig. 5. **Special case for the first overshoot magnitude where first overshoot is finite**

will call for a maximum magnitude of controller output (actually an algebraic minimum). Point a indicates the value of x required, as thus established and, hence is the starting point of the transient associated with this test. In test 2, an external signal is applied to the system so that minimum controller output is obtained in the steady state. Thus $\dot{x}=0$, but x takes on some positive value needed for the controller output to be minimum. During test 2 the trajectory starts at point a and proceeds to the origin along the trajectory of the linear system. The transient of this trajectory is the test-2 result, and the duration of this transient is the forcing-utilization contribution. In Fig. 7, the basic trajectory ba is constructed by displacing the optimum trajectory qo horizontally a distance ao.

In this discussion, only those systems are considered for which the large-step response has an overshoot with a single-mode controller. An overshoot of magnitude greater than the distance oa is considered as not permissible after the second controller mode is introduced. A point d is plotted in Fig. 7, from the data of test 3, with od the magnitude of the first overshoot.

Consider the transient associated with the trajectory that moves up to q in an optimum manner, then from q to b horizontally, down the curve ba to point a, and then from a into the origin along the trajectory of the linear system. The duration of this transient is simply contribution 1, plus contribution 2, plus the time required to move from q to b. The time to move up to q is the part of contribution 1 for maximum controller output; from b to a, the part of contribution 1 for minimum controller output; and from a to the origin is the second contribution. The time associated with the transition from q to b is very small compared with the time from a to o, since, in the first case $\dot{x}$ is very large; in the second, small.

Fig. 8 shows the controller output for the transient with the base trajectory described in the foregoing, and Fig. 9 shows the actual transient. For this transient

$$\text{Contribution 1} + \text{contribution 2} = \text{actual duration} \qquad (1)$$

considering time $t_b - t_q$ as negligible.

The trajectory of the actual response with only a single-mode controller does not follow this base trajectory, since it has a value $x=od$ for $\dot{x}=0$. It is informative to examine the conditions which must exist for the actual trajectory to lie above the base trajectory. For this case, there exists a region above the base trajectory in which the system is linear in the sense that the controller output is not minimum. The curve ba is a trajectory for minimum controller output. Thus, unless a region exists above ba in which the controller output is not minimum, it must follow that the actual trajectory for the system transient will always lie below the base trajectory ba. When a second mode of operation for the controller is designed, its purpose is to remove this region above the base trajectory.

For example, if the controller second mode is such as to make the boundary fea (Fig. 7) the boundary for minimum controller output (when the actual trajectory is above fea, the controller output is minimum), then the actual trajectory will have a path below fe and will not intersect the base trajectory ba until the point e. The duration of the transient for this trajectory is greater than the sum of contribution 1 plus contribution 2; however, this transient is of little interest, since it is not obtained for the optimum second mode of operation.

If, on the other hand, the boundary for minimum controller output consists of line bea, then the actual trajectory is approximately shown by the dotted lines. The difference in duration of the transient following the dotted path and the path qbe is small compared with the time required from the point a to the origin, since in one case the $\dot{x}$ is large and in the other case it is small. Therefore, with the boundary bea for minimum controller output, the duration of the actual transient is approximately equal to the sum of the first and second contributions.

The magnitude of the first overshoot for the large-step response with only a single-mode controller is a measure of the need for a second mode. For example, if with a single-mode controller only a small overshoot is obtained, it is possible to obtain only slight improvement by instrumenting a second mode. On the other hand a large overshoot indicates that much improvement is possible. Thus the entire large-step response transient is not required to establish this, but only the first overshoot magnitude.

In the process of instrumenting a second mode this first overshoot magnitude also provides a guide for the intrumentation. For a properly designed second mode only a slight overshoot will be obtained. There are several additional practical ramifications involved in this second-mode design. First, having achieved the condition of small overshoot for a given large step does not guarantee that this second mode is necessarily the optimum for this particular step nor does it guarantee that only a small overshoot will be obtained for other magnitudes of large steps. Furthermore, it may be desirable to tolerate some overshoot in the response to the large-step function in order to achieve a wider range for the linear mode of the controller. It may also be possible to modify the design of the single mode so as to reduce overshoot for the large step. In doing this it is very likely that some small signal performance will be sacrificed, but this compromise may overweigh the disadvantages of having to instrument for dual mode operation.

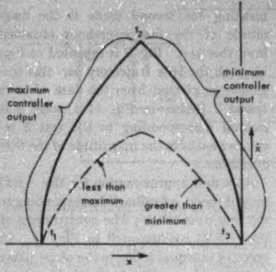
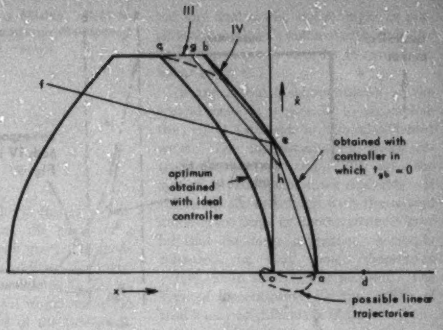


Fig. 6. System output trajectory for optimum large-step response

Fig. 7 (right). Construction to demonstrate that contribution 1 + contribution 2 = actual duration



III
IV
c
g
b
f
$\dot{x}$
e
obtained with controller in which $t_{gb} = 0$
optimum obtained with ideal controller
h
o
a
d
possible linear trajectories
x

The foregoing considerations are associated with instrumenting the second mode and are not important at the evaluation stage. In summary, a large overshoot with only a single-mode controller indicates that improvement would be obtained using a dual mode. This overshoot may also be used as a guide for instrumenting the second mode, but there are additional considerations, as mentioned.

In Fig. 7, if a boundary line *gha* were obtained where *gh* was slightly below *be*, the actual transient would still have an overshoot; but in this case the magnitude of the overshoot would be less than *oa*, and the trajectory of the actual response lies between the base trajectory and the optimum trajectory. In this event the duration of the actual transient would be less than the sum of contributions 1 and 2. It is possible to obtain only a small improvement, however; and if the boundary *gh* is too much below *be*, the actual duration will be greater than the sum of contributions 1 and 2.

In general, then, with the optimum second mode the following are predicted:

Actual duration of response = (contribution 1) + (contribution 2) (2)

Magnitude of the actual overshoot = *oa* (3)

From a simple open-loop test for contribution 1 and a simple closed-loop linear test for contribution 2, the actual duration and overshoot of the optimum large-step transient with a dual-mode controller when the controller dynamics are not neglected are predicted before the second mode is actually instrumented.

Equations 2 and 3 are exact in the limiting case for contribution 2 equal to zero and *oa* equal to zero. If the contribution 2 is large compared with contribution 1 it is usually possible to obtain an optimum large-step response which is less than the sum of contributions 1 and 2. This fact does not detract from the usefulness of these equations, since it is obvious when contribution 2 is large compared with contribution 1 that the forcing ability is not being utilized. In this case an improvement in large-step response can usually be made by simply lowering the saturation level of the controller. A better solution may be to shift power gain from the controller to the output device thus improving the dynamic properties of the controller and reducing the second contribution.

The three desired items of information previously discussed are obtained from the three proposed evaluation tests. These are

1. The large-step response of the system with only a single-mode controller. In test 3 the first overshoot magnitude of this response has been determined. It has been shown, however, that the magnitude of this overshoot constitutes sufficient data to determine whether the instrumentation of a second mode for the controller is warranted.
2. The optimum large-step response of the system with a dual-mode controller not neglecting the dynamic properties of the controller. This datum is obtained directly as the sum of the forcing-ability contribution of test 1 and the forcing-utilization contribution of test 2.
3. The optimum large-step response of the system with a dual-mode controller neglecting the dynamic properties of the controller. This datum is the forcing ability contribution of test 1.

The difference between data 2 and 3 is due to the dynamics of the controller and is simply the forcing-utilization contribution of test 2. The forcing-utilization contribution, however, has meaning only as it is compared to the forcing-ability contribution.

If the optimum large-step response neglecting controller dynamics of test 1 is not satisfactory, then it is immediately apparent that the output device or the controller forcing must be modified. The result of test 1 represents the ultimate in large-step response and is hence the criterion in the instrumenting of a second mode or in the improving of controller dynamics.

Performance of the Forcing-Ability Test 1

Only test 1 requires special discussion. The performance of tests 2 and 3 are obvious from their description. The discussion of this section pertains only to that class of second-order output devices for which optimum step response is obtained with one cycle of maximum controller output and one cycle of minimum controller output although the method can be extended to other systems as will be discussed.

The concept for performing test 1 is the same whether a computer or an analytical technique is used. At the instant of application of the large-step command or disturbance, the controller output should be simultaneously maximum. Thus, at the same instant these two conditions are introduced to start the calculation of the open-loop transient. Initially, x is negative and $\dot{x}=0$; the final condition is $x=0$ and $\dot{x}=0$. At some instant the controller output is reversed to minimum output for the duration of the transient. With minimum output from the controller, $\dot{x}$, which is positive during the period when maximum output is applied, will not only go to zero but will then become negative. This is of no concern, since the transient is assumed completed when $\dot{x}$ becomes zero as it passes from a positive to a negative value.

The only detail missing is the desired

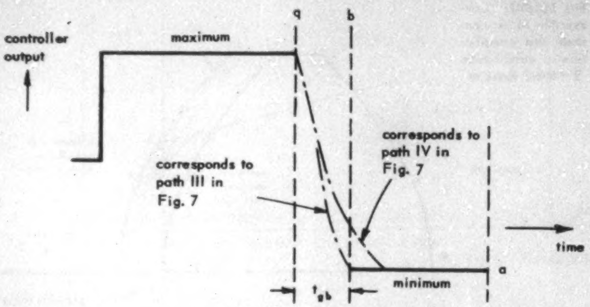

q
b
controller output
maximum
corresponds to path III in Fig. 7
corresponds to path IV in Fig. 7
time
a
minimum
t_{gb}

Fig. 8. Controller output for the base trajectory

time of reversal from maximum controller output to minimum controller output. One simple way of determining this instant is by cut-and-try. Using a computer, an instant is chosen and the computer is put on "hold." The controller output is reversed to minimum and the transient is resumed. The operation is identical analytically. If the trial reversal is made too soon, $\dot{x}$ will equal zero when x is still negative; if too late, x will have become positive when $\dot{x}=0$. The value of x when $\dot{x}=0$ is a good approximation to the error in reversal; thus, in two or three trials the result that x becomes zero and $\dot{x}$ becomes zero simultaneously is achieved. The duration of this transient is the desired result. With this approach, the result obtained is a plot of x as a function of time. There is no difficulty in determining when $\dot{x}=0$. The proper switching instant is also available in certain cases from a direct calculation. In most cases, however—for example, when the output device contains several nonlinearities—the cut-and-try procedure is easier. When the desired transient is obtained, it is important to record not only the output signal but also the signal behavior at all nonlinearities.

A second way of determining the proper reversal instant which is more informative than cut-and-try, is to record x versus $\dot{x}$ in the phase plane rather than x versus time. The technique is to record x versus $\dot{x}$ for a period well past the optimum reversal instant with maximum controller output. Then the system is set up in the final steady state with minimum controller-output signal for negative time, and again x versus $\dot{x}$ is recorded. Negative time is obtained by replacing the operator s in the signal-flow diagram by $-s$. The intersection of these two curves is the optimum reversal point. When the application of the step does not change the transfer function of the output device, the second trace will also be the optimum reversal curve for all amplitudes of steps.

Instrumenting the Second Mode

It was shown in Fig. 7 that an overshoot in the large-step response with only a single-mode controller is due to the existence of a region above the base trajectory with nonminimum controller output. The second mode may be designed by feeding back a signal $-f(\dot{x})$ into the controller to remove this region above the base trajectory. It is important, however, not to feed back more $-f(\dot{x})$ than is necessary, since the transient duration would then be made unnecessarily long. Since it may be desirable to perform test 3 on the actual hardware, it is assumed that the only datum available for instrumenting the second mode is the magnitude of the first overshoot obtained from this test. Fig. 7 is repeated as Fig. 11, with the base trajectory *ba;* this trajectory is plotted from the data of tests 1 and 2. Also, in Fig. 11 a point d is located corresponding to the test 3 result, where od is the magnitude of the first overshoot.

As a first approximation for the $-f(\dot{x})$ second mode, a feedback corresponding to the line oc is tried. The construction of this $-f(\dot{x})$ is illustrated in Fig. 11. In order to interpret a line in the phase plane in terms of a $-f(\dot{x})$ feedback in a feedback-control system the equation is

$$-x-f(\dot{x})=0$$

where x is the value corresponding to the $\dot{x}$ on the line. This arrangement is further illustrated by Fig. 10 with $x_{\text{reference}}=0$. With this second mode instrumented, test 3 is performed again and the magnitude of the first overshoot recorded. It is important that this overshoot be at least of magnitude oa or a little more. If an overshoot of oa or a little more is not obtained, a second trial $-f(\dot{x})$ is made, with the point c moved to the right if an overshoot less than oa is obtained, and moved to the left if an overshoot greater than oa is obtained. When, after several trials, an overshoot of magnitude oa or slightly more is obtained, the second trial $-f(\dot{x})$ is made, as shown in Fig. 11. After instrumenting the second trial $-f(\dot{x})$, test 3 is again performed to obtain an overshoot of oa or slightly more. If the overshoot is more than oa, point f is moved down; if less, point f is moved up. When an overshoot of oa or slightly more is obtained, the duration of the actual tran-

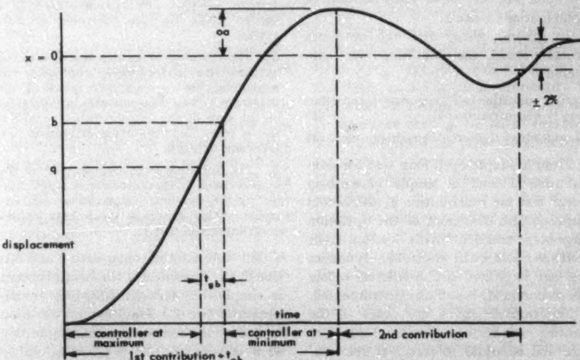


Fig. 9. Transient of the base trajectory

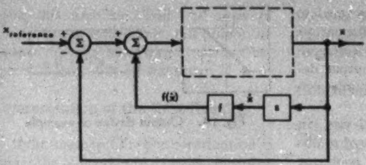


Fig. 10. Mixing $-f(\dot{x})$ to obtain the best possible second mode

sient is approximately equal to the sum of contributions 1 and 2.

It is desirable to try to obtain an actual duration less than the sum of contributions 1 and 2. This is accomplished by making a third trial, as shown in Fig. 11. The third trial is for a line $c''f''$ parallel and below $c'f'$. In this case, the $c''f''$ is chosen to minimize the time duration of the actual transient. As $c''f''$ is moved downward from $c'f'$, the actual duration will become shorter, will reach a minimum, and then will start becoming longer again.

The reasons for making trials 1, 2, and 3 as indicated are as follows:

Since the overshoot *od* is a distance *ad* to the right of the base trajectory *ba*, it is assumed as a first approximation that, with only a single mode, a trajectory displaced *ad* from the base trajectory is the actual path. It is further assumed that the single-mode boundary for minimum controller output is the line *ga*, and that the region *bgh* is the region above the base trajectory in which the controller output is not minimum. The first trial $-f(\dot{x})$ is intended to move the boundary from *gha* to the line *ba*. It is important, however, that not too much $-f(\dot{x})$ be used; otherwise, some trajectory below the base trajectory *ba* would be used. Unless the overshoot with the first trial $-f(\dot{x})$ is *oa* or greater, there is no way of knowing how far below the base trajectory the actual trajectory is, since for an overshoot less than *oa* the controller output is not minimum when $\dot{x}=0$, and therefore there is no way of telling how long the controller output has been not minimum. If it has been not minimum for a large part of the actual trajectory, then the duration of the actual transient could be greater than the sum of the first and second contributions. Thus it is necessary to trim trial 1 so that an overshoot slightly more than *oa* is obtained.

With the first trial $-f(\dot{x})$, the boundary for minimum controller output *ba* creates a large region between line *ba* and trajectory *ba* of minimum controller output. This region is unnecessary and also undesirable. The point *f* for the second mode should be as high as possible in order that the first mode shall have the largest range possible: For example, if this is a positional servomechanism, the error for a velocity output is much greater in mode 2 than in mode 1. In addition, for intermediate-size steps, the line *ba* is the reversal line, and reversal would occur too soon and the duration of the intermediate-size step would be unnecessarily long. The second trial is then intended to adjust the $-f(\dot{x})$ feedback so that only the region *bgh* is removed.

After the correct c' and f' have been determined, it is then possible to trim to $c''f''$, and the transient of minimum duration is obtained. If $c''f''$ is too far below $c'f'$, however, the transient may again have a duration greater than the sum of the first and second contributions.

In some cases the point *f* at which the second mode starts may be too low, and the linear or first-mode range is too narrow to obtain good linear preformance. If the point *f* is raised, an overshoot is obtained: for example, a point f''' would result in an overshoot of approximate magnitude *jk*, as shown in Fig. 11. This overshoot is less than the overshoot *od* without any dual mode, and it might be possible to permit this much overshoot in the large-step response in order to extend the linear range to f'''.

In the actual instrumentation of the second mode, it is possible to accomplish the $-f(\dot{x})$ feedback in several different ways. The instrumentation which introduces the least complexity into the system is usually the most desirable. If the hardware associated with the second mode is too large or cumbersome, it may be that the large overshoot, is simply tolerated, or that some compromise $-f(\dot{x})$, as in trial 1, will be used. The type of discontinuity for the $-f(\dot{x})$ of trial 2 may be difficult to obtain.

The most straightforward manner would be to feedback $\dot{x}$ through a nonlinearity, as shown in Fig. 12. This can be accomplished by adding the two signals shown in Fig. 12. One signal is a negative feedback and the other is a limited positive feedback. A limited positive feedback may even be inherent in the system; and rather than eliminate it as undesirable, it may be possible to instrument it to obtain the desired second mode.

If the system is required to have good large-step response over a wide range of large steps, the final result is a many-moded system only in the sense that the $-f(\dot{x})$ is a nonlinear function $\dot{x}$. This nonlinear function is easily approximated by performing the procedure of this paper for several sizes of large steps.

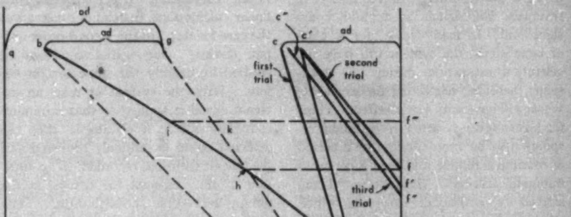


Fig. 11. Construction to instrument the second mode

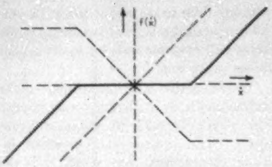


Fig. 12. Generation of $f(\dot{x})$

Systems with Higher-Order Output Devices

Up to this point only those output devices have been considered for which the optimum large-step response is obtained with one reversal of the controller-output signal. Not only are these devices easier to understand conceptually but also the mechanics of instrumenting for a single reversal are easier than for several reversals.

In theory, at least, tests 1 and 2 are defined in such a way that the application to systems with higher-order output devices would require no additional modification of the definition. Test 1 involves finding the optimum large-step response of the output device, assuming that the controller-output signal can be varied between the maximum and minimum level instantaneously. Thus there is nothing in the definition itself which would prevent more than one reversal. Test 2 does not restrict the order of the end device either.

For example: if the output device is of a third order and requires two reversals, the signals x, $\dot{x}$, and $\ddot{x}$ are significant. In many cases, for a class of large steps, the system will have acceleration saturation during the transient; therefore, neglecting the second reversal will have only a small effect. Thus, for higher-order systems the optimum response may be approximated by a period of maximum forcing and then a period of minimum forcing or the inverse. Then, instead of additional reversals, which would give optimum response, the system returns to the linear mode to complete the transient. First consider the exact case. Here test 1 is performed to obtain the optimum response. In general, the number of reversals required is equal to $(n-1)$ where n is the order of the output device. The exception to this is demonstrated by the second-order output device, which is linearly a second-order underdamped quadratic.

For a third-order output device, in general, two reversals are required to obtain optimum response. The points of these reversals can be determined as a function of x, $\dot{x}$, and $\ddot{x}$. Instead, an approximate procedure for looking at the higher-order output device, which simplifies the calculations and (most important) the instrumentation, is usually to be preferred. In this case, regardless of the order of the output device, it is assumed that the controller optimum output is a period of maximum output followed by a period of minimum output or the inverse and that during the remainder of the transient, the system performs linearly. In this case, test 1 is performed by cut-and-try. First, with the system in the initial steady state, maximum forcing is applied simultaneously with the application of the large step, recording x as a function of time. Then, by cut-and-try, the point is determined so that applying minimum forcing from this point on, results in x and $\dot{x}=0$ simultaneously; however, since this is a higher-order output device, $\ddot{x}$ is not necessarily equal to zero. In other words, since $x=0$ and $\dot{x}=0$, the transient would be completed except that $\ddot{x}\neq 0$, and the controller output is not at its final steady value, but is at minimum value. Thus there are two initial conditions to be studied using linear calculations, instead of one, as was the case for test 2 in the second-order output device. One initial condition is studied by exactly the same test as before. With the system at rest, an external signal is applied so that minimum controller output is obtained; then the external signal is suddenly removed and the output transient recorded. The duration of the transient, as before, is the forcing-utilization contribution. The second initial condition is studied by inserting an external signal so as to make $\ddot{x}$ appear to have the value noted when x and $\dot{x}=0$ at the end of the optimum response. The mechanics of introducing

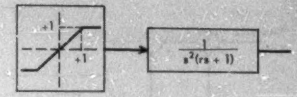


Fig. 14. Output device of example

this external signal are shown in Fig. 13. The external signal is suddenly removed and the output transient is recorded. The duration of this transient is the $\ddot{x}_d$ contribution. Fig. 14 illustrates a possible higher-order output device and controller maximum and minimum output. Fig. 15 shows the methods for introducing the external test signals to find the second and $\ddot{x}_d$ contributions.

This procedure is repeated for $\dddot{x}_d$, etc., for systems of order higher than three. It might appear that these initial conditions should be applied simultaneously, since this is in fact the way in which the system performs during the large-step transient.

The system, however, may be represented by a linear model for these transients, so that superposition applies and the transients for each initial condition may be investigated singly if desired. By investigating these transients singly, their durations may be compared, and the relative magnitudes of these durations are significant. The duration of the optimum large-step response with a dual mode, not neglecting the controller dynamics, is predicted to be the sum of the forcing-ability contribution and the larger of the forcing utilization contribution or the $\ddot{x}_d$ contribution. Thus, if the forcing-utilization contribution is larger than the $\ddot{x}_d$ contribution, very little is lost by not instrumenting for the extra reversals. This conclusion may seem trivial, in that it amounts to saying that if the output device behaves like a second-order output device even though it is of higher order, then the procedures for second-order devices may be used. The significant part is, that the $\ddot{x}_d$ contribution is a measure of how nearly the higher-order system behaves like a second-order system. Those systems for which the actual behavior of the output device is higher than second order are not treated in this paper; how-

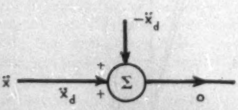


Fig. 13. Arrangement for determining the $\ddot{x}_d$ contribution

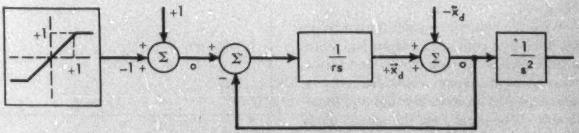


Fig. 15. Placing external signals to find the second and $\ddot{x}_d$ contributions

ever, this does not limit the order of the system to second order but simply the output device. Most output devices behave, at least, like second order.

Interpretation of the Test Results

If the forcing-utilization contribution is large, the indication is that the controller forcing may be reduced without sacrificing large-step response. A better solution may be to shift power gain from the controller to the output device thus improving the dynamic characteristics of the controller. There is a practical limit to this shifting, since an increased power gain for the output device will usually mean less ability to be forced. Thus a compromise is indicated and this compromise is facilitated by comparing the forcing-ability contribution and the forcing-utilization contribution.

If the first overshoot magnitude of the large-step response with a single-mode controller is large, it is indicated that improvement in the large-step response is possible by instrumenting a second mode. If for a higher-order output device the $\ddot{x}_d$, etc., contributions are small, the indication is that, despite the higher order, the dual-mode controller need only produce one reversal in order to obtain good large-step response.

Conclusions

A technique has been presented for evaluating a feedback-control system for a large-step test function. This technique is intended to complement the study of a system for small signal variations. The type of test signal to use for small variations will depend on the application and may be periodic, aperiodic, or statistical. In general, the large step is a good test function for large signal variations since, in an error-minimizing system, it is usually only for sudden commands or disturbances that a large error will occur.

An exact study of the large signal variations in a system is difficult due to nonlinearities; this is particularly the case if the system is of higher order. In addition, the representations of systems for large signal variations is usually very approximate, thus an exact study is not usually warranted. Instead, an approximate technique which will give useful information from simpler calculations is suggested.

Using the techniques of this paper it is possible to obtain information regarding the improvement to be obtained by instrumenting a dual-mode controller for a system which has, initially, only a single-mode controller and to obtain information regarding saturation levels and power gains for cascaded components. If it is decided to modify saturation levels or power gains the tests will provide a guide for this modification. If it is decided to instrument a second mode of operation for the controller the tests will provide a guide for this instrumentation.

As a possible extension to the work of this paper the possibility of reducing the approximate nature of the results without increasing the complexity of the calculations is suggested.

References

1. SERVOS WITH TORQUE SATURATION, W. Hurewicz. *Radiation Laboratory Report 592*, Massachusetts Institute of Technology, Cambridge, Mass., 1944.

2. NONLINEAR TECHNIQUES FOR IMPROVING SERVO PERFORMANCE, D. C. McDonald. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 400–21.

3. A PHASE-PLANE APPROACH TO THE COMPENSATION OF SATURATING SERVOMECHANISMS, Arthur M. Hopkins. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

Noninteracting Controls in Linear Multivariable Systems

R. J. KAVANAGH
STUDENT MEMBER AIEE

Synopsis: In linear multivariable control systems, a change in any input variable will usually result in changes in all output variables. It is useful for certain applications to obtain a system in which such interaction between controls does not occur. This paper uses previously derived matrix methods to synthesize a system with noninteracting controls by applying feedback around the given multivariable system. In addition to providing simple design equations for such systems, equations are also derived which give the input-output response and the disturbance response of the synthesized systems. It is shown that the matrix approach leads to a more simple and useful solution to the noninteraction problem than was the case in a previous approach to this problem.

THIS paper is concerned with the class of systems known as linear multivariable control systems and with one of the results of applying feedback around such systems. Multivariable systems are those systems in which one or more output variables are controlled by one or more input variables. As in the field of single-variable control systems, the constraint of linearity has been imposed as a necessary step before the problem of nonlinear systems can be considered. The class of linear multivariable systems may be seen to include the common linear single-variable system as a special case.

In general, multivariable systems having more output variables than input variables are feasible. However, such systems are not useful in practice and will therefore be omitted in this study. A block diagram of a multivariable system is shown in Fig. 1. This is a system in which there are n inputs and m outputs, where $n \geq m$.

Multivariable control systems occur in practice in various fields. The control of a steel rolling mill is one example[1] and of turbojet engines is another.[2] The linearity of such systems is open to question but with a restricted range of operation about a fixed point, the linear theory may be applied with useful results.

The Noninteraction Problem

An important problem in the study of multivariable systems is that of synthesizing a system which has noninteracting controls. Essentially, this problem consists of applying some form of feedback around a system such that in the resulting system, variation of each input quantity will cause only one output variable to change, and in addition, each output is

---

Paper 57-41, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 22, 1956; made available for printing November 16, 1956.

R. J. KAVANAGH is with the University of Toronto, Toronto, Ont., Canada.

The author acknowledges the encouragement of Dr. J. M. Ham, University of Toronto. Acknowledgement is also due the National Research Council of Canada for permission to publish this paper and for financial assistance received during the period when this work was performed.

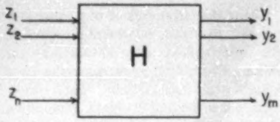


Fig. 1. The general linear multivariable system

affected by changes in only one input variable. It is apparent that for the case of a square system having an equal number of inputs and outputs a one-to-one relationship occurs between the system inputs and outputs. In the case of a nonsquare system however, it is apparent that such a relationship is not possible. Therefore, for the nonsquare case, noninteraction is defined such that the application of feedback around a system results in a one-to-one relationship between the output variables and an equal number of over-all system inputs. In addition, the remaining over-all system inputs are used to control an equal number of the basic system inputs such that a one-to-one correspondence prevails and such that the system outputs are unaffected by changes in these variables. It may be seen that the square case is a particular case of the general nonsquare system.

The condition of noninteraction is often desirable in multivariable control systems. It removes cross-coupling between variables thus permitting any output quantity to be adjusted independently of other output quantities. It is useful, for example, in the field of turbopropellor engines where the system inputs are the propellor blade angle and the fuel rate. The system outputs are the speed of rotation and the turbine-inlet temperature. Noninteracting controls could therefore enable the speed of the engine to be controlled only by the fuel rate. In addition the fuel rate controls only the engine speed. It may be seen that the condition of noninteraction would be of great value in other multivariable control systems such as steel rolling mills.

METHOD OF SOLUTION

In 1950, Boksenbom and Hood[2,3] considered the problem of noninteraction in multivariable systems and obtained a solution which enabled systems to be modified by feedback to give noninteraction as defined. Their method employed to a certain extent the concept of the matrix. Since that time Cruickshank[4] in England formulated matrix equations which may be used to describe multivariable systems and independently, the author developed a more general and complete method of applying matrix algebra to the description of multivariable systems.[5]

In this paper, a solution to the noninteraction problem will be given which employs the matrix approach to multivariable systems described by the author in a previous paper.[6] It will be found that the results obtained by Boksenbom and Hood are verified but in a more compact and useful form. Some additional results are also derived. Besides yielding an alternative derivation to that obtained by Boksenbom and Hood, this paper also serves as an illustration of the use of matrix methods in the analysis and synthesis of multivariable control systems.

Basic Theory of Linear Multivariable Systems

The linear multivariable system shown in Fig. 1 may be suitably described by a set of simultaneous linear equations using the concept of the Laplace transformation,

$$\begin{aligned} y_1 &= h_{11}z_1 + h_{12}z_2 + \ldots + h_{1n}z_n \\ y_2 &= h_{21}z_1 + h_{22}z_2 + \ldots + h_{2n}z_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= h_{m1}z_1 + h_{m2}z_2 + \ldots + h_{mn}z_n \end{aligned} \qquad (1)$$

These equations describe the relationship between the outputs, y_1, y_2, ... y_m and the inputs x_1, x_2, ... x_n. The h's are analogous to the normal transfer function in the case of single-variable systems.

Equation 1 may now be written in matrix form,

$$Y = HZ \qquad (2)$$

where Y is an $m \times 1$ column matrix, Z is an $n \times 1$ column matrix and H is the $m \times n$ matrix,

$$H = \begin{vmatrix} h_{11}h_{12}\ldots h_{1n} \\ h_{21}h_{22}\ldots h_{2n} \\ \ldots\ldots\ldots \\ h_{m1}h_{m2}\ldots h_{mn} \end{vmatrix}$$

This matrix, called the transfer matrix, whose elements are functions of the complex frequency variable, completely describes the given multivariable system.

There are many ways in which additional signal paths or circuit elements may be added to the system of Fig. 1 in order to obtain a new over-all system with changed characteristics. For the purpose of this paper, two restrictions on such modifications will be made, as follows:

1. No additional circuit elements may be placed in the output paths of the basic system.
2. The resultant over-all system must still be linear.

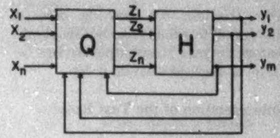


Fig. 2. Linear multivariable system with feedback computer

Restriction 1 is imposed because, in general, a difference in power level occurs between the input and output of systems. The work of this paper is concerned with obtaining changed system characteristics by operating upon signals at low power level. This situation is analogous to that prevalent in the design of single-variable systems. With this restriction in mind, it may be seen that the most general method of modifying the given system is that shown in Fig. 2. Q is a device, which may be called a computer, such that the inputs to the original system, z, are now some function of the over-all system inputs, x, and the system outputs, y. It now follows from restriction 2 that Q must be a linear device such that the z's are actually linear combinations of the x's and the y's.

The computer Q may now be described by the following set of linear equations by use of the Laplace transformation,

$$\begin{aligned} z_1 &= c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n + k_{11}y_1 + \\ &\qquad k_{12}y_2 + \ldots + k_{1m}y_m \\ z_2 &= c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n + k_{21}y_1 + \\ &\qquad k_{22}y_2 + \ldots + k_{2m}y_m \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_n &= c_{n1}x_1 + c_{n2}x_2 + \ldots + c_{nn}x_n + k_{n1}y_1 + \\ &\qquad k_{n2}y_2 + \ldots + k_{nm}y_m \end{aligned} \qquad (3)$$

These equations may be written in matrix form,

$$Z = CX + KY \qquad (4)$$

where X is an $n \times 1$ column matrix and C and K are given by

$$C = \begin{vmatrix} c_{11}c_{12}\ldots c_{1n} \\ c_{21}c_{22}\ldots c_{2n} \\ \ldots\ldots\ldots \\ c_{n1}c_{n2}\ldots c_{nn} \end{vmatrix}$$

$$K = \begin{vmatrix} k_{11}k_{12}\ldots k_{1m} \\ k_{21}k_{22}\ldots k_{2m} \\ \ldots\ldots\ldots \\ k_{n1}k_{n2}\ldots k_{nm} \end{vmatrix}$$

These two transfer matrices completely describe the computer.

To determine the over-all characteristics of the system shown in Fig. 2, it is necessary to eliminate Z by combining equations 2 and 4. There results the relationship

$$Y = (I - HK)^{-1}HCX \qquad (5)$$

where I is the $n \times n$ identity matrix. This equation gives, in matrix form, the relationship between the input and output variables for the over-all system. Equation 5 may be written in shorter form as the following:

$$Y=H'X \tag{6}$$

where H' is the transfer matrix for the over-all system. Hence there results the relationship which is given by the following:

$$H'=(I-HK)^{-1}HC \tag{7}$$

or

$$H'=HC+HKH' \tag{8}$$

Equations 7 and 8 are important results. They are two forms of the same equation relating the transfer matrices of the multivariable system with and without the computer. It may be seen that equation 7 is a generalization of the equation relating the open and closed loop-transfer functions of a single-variable control system.

Equation 8 represents a set of mn simultaneous equations since each term of the equation is an $m\times n$ matrix. For future purposes it will be useful to be able to write out these mn equations explicitly. It may be seen that the typical equation is given by

$$h'_{ij}=\sum_{q=1}^{n}h_{iq}c_{qj}+\sum_{p=1}^{m}\sum_{q=1}^{n}h_{iq}k_{qp}h'_{pj} \tag{9}$$

Solution of Noninteraction Problem

The form of system to be studied is shown in Fig. 3. The basic system H is shown as having five inputs and three outputs as a specific example of the general system having n inputs and m outputs where $n>m$.

It is desired to modify the foregoing system by feedback such that a one-to-one relationship will occur between the three outputs r_1, r_2, and r_3, and three of the over-all system inputs R_1, R_2, and R_3. The two remaining over-all system inputs S_4 and S_5 will control the two basic system inputs s_4 and s_5 such that a one-to-one relationship prevails. Also shown in Fig. 3 are five disturbance inputs, V_1, V_2, V_3, V_4, and V_5. These are assumed to act upon the inputs to H in an additive manner.

In this particular case where a controlled variable is required to respond to a certain input variable, it seems reasonable to select simple subtractive devices for the feedback comparators. The resulting error signals are then applied to the computer Q. The problem to be solved is therefore the specification of the elements of Q such that the required system behavior is obtained.

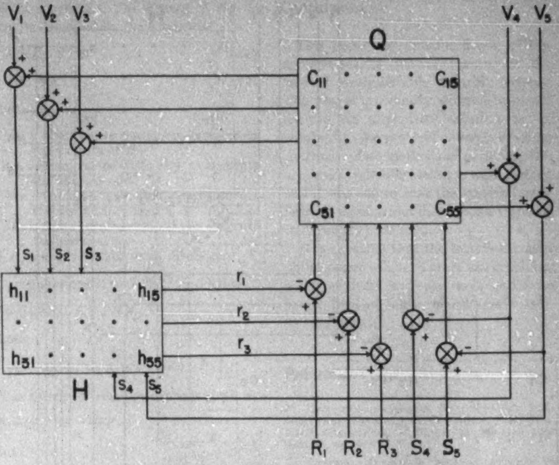


Fig. 3. Control configuration for obtaining noninteraction

Control servomechanisms and measuring instruments have been omitted from Fig. 3. In the case where such elements are present in addition to H and Q, their effect may be taken into account by suitably modifying the elements of H and Q. Hence the following equations will apply in all cases.

The first step in the solution of the problem is to redraw the system in such a way that the applicability of the fundamental matrix equation becomes obvious. Consider the system shown in Fig. 4. It may be seen that the system represented in this figure is identical to that shown in Fig. 3. The system is now recognizable, however, as a particular case of the general multivariable feedback system in which the controlled system has a square matrix. By comparing the system of Fig. 4 with the general system, it follows that

$$H=\begin{vmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} \\ h_{21} & h_{22} & h_{23} & h_{24} & h_{25} \\ h_{31} & h_{32} & h_{33} & h_{34} & h_{35} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$$C=\begin{vmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} \end{vmatrix}$$

$$K=-C$$

The noninteraction condition means that the over-all system matrix should have the following form:

$$H'=\begin{vmatrix} h_{11}' & 0 & 0 & 0 & 0 \\ 0 & h_{22}' & 0 & 0 & 0 \\ 0 & 0 & h_{33}' & 0 & 0 \\ 0 & 0 & 0 & h_{44}' & 0 \\ 0 & 0 & 0 & 0 & h_{55}' \end{vmatrix}$$

These quantities are related by equations 7 and 8. Substitition for K in equation 8 gives, for this case,

$$H'=HC-HCH' \tag{10}$$

The synthesis problem now becomes one of solving equation 10 for C. It may be seen that the formal solution for C is given by

$$C=H^{-1}H'(I-H')^{-1} \tag{11}$$

Thus, if H', the response of the noninteracting system, is given, then a matrix C is yielded by equation 11 such that the system will have noninteracting controls as well as the desired response. It is of interest to note that H' cannot become the identity matrix since equation 11 has no solution in this case.

An alternative solution of equation 10 may be obtained by writing it in an expanded form using equation 9, in the following manner:

$$h'_{ij}=\sum_{k=1}^{n}h_{ik}c_{kj}-\sum_{p=1}^{n}\sum_{k=1}^{n}h'_{pj}h_{ik}c_{kp} \tag{12}$$

for an n input system. For a system with noninteraction, h'_{ij} is zero for $i\neq j$.

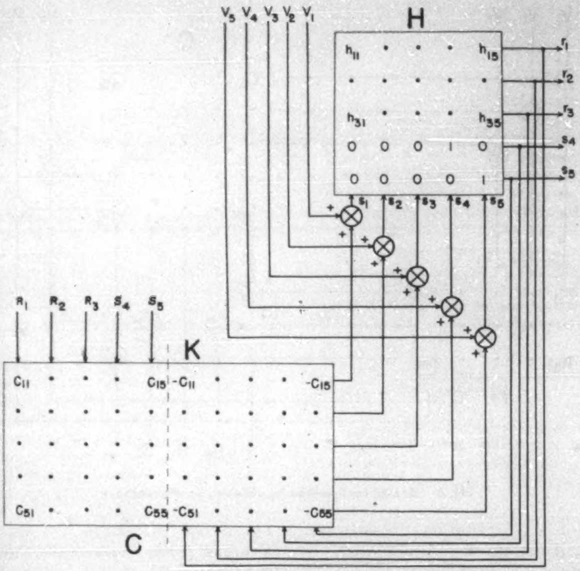


Fig. 4. Redrawn configuration for obtaining noninteraction

Hence in this case equation 12 becomes

$$0=\sum_{k=1}^{n} h_{ik}c_{kj}-h'_{jj}\sum_{k=1}^{n} h_{ik}c_{kj} \tag{13}$$

where $i=1, 2, \ldots\ldots, j-1, j+1, \ldots\ldots n$. Or, assuming $h'_{jj}\neq 1$,

$$\sum_{k=1}^{n} h_{ik}c_{kj}=0 \tag{14}$$

where $i=1, 2, \ldots\ldots, j-1, j+1, \ldots\ldots n$.

For any given i, equation 14 represents $n-1$ equations in n unknowns, c_{kj}. This enables the ratio of any two elements in each column of C to be determined.

For any given column of C, equation 14 may be written in expanded form as:

$$\begin{aligned} &h_{11}c_{1j}+h_{12}c_{2j}+\ldots\ldots+h_{1n}c_{nj}=0\\ &h_{21}c_{1j}+h_{22}c_{2j}+\ldots\ldots+h_{2n}c_{nj}=0\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &h_{n1}c_{1j}+h_{n2}c_{2j}+\ldots\ldots+h_{nn}c_{nj}=0 \end{aligned} \tag{15}$$

with the jth row omitted. Dividing each of the equations throughout by c_{sj} and moving the coefficient of c_{sj} to the right-hand side in each case,

$$\begin{aligned} &h_{11}\frac{c_{1j}}{c_{sj}}+h_{12}\frac{c_{2j}}{c_{sj}}+\ldots+h_{1n}\frac{c_{nj}}{c_{sj}}=-h_{1s}\\ &h_{21}\frac{c_{1j}}{c_{sj}}+h_{22}\frac{c_{2j}}{c_{sj}}+\ldots+h_{2n}\frac{c_{nj}}{c_{sj}}=-h_{2s}\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &h_{n1}\frac{c_{1j}}{c_{sj}}+h_{n2}\frac{c_{2j}}{c_{sj}}+\ldots+h_{nn}\frac{c_{nj}}{c_{sj}}=-h_{ns} \end{aligned} \tag{16}$$

with the jth row and the sth column on the left-hand side omitted. Equations 16 are a set of $n-1$ equations in the $n-1$ unknowns, c_{rj}/c_{sj} for $r=1, 2, \ldots\, s-1, s+1, \ldots\, n$. The general solution of these equations is given by

$$\frac{c_{rj}}{c_{sj}}=\frac{|H_{jr}|}{|H_{js}|} \tag{17}$$

where $|H_{jr}|$ and $|H_{js}|$ are the cofactors of h_{jr} and h_{js} respectively.

Equation 17 enables the ratio of any two elements in the same column of the computer to be determined from a knowledge of the given system matrix. Consequently, if one element in each column of the computer matrix is known, all of the remaining computer elements may be determined by means of equation 17. It should, of course, be remembered that the solution of equation 16 by the determinant method becomes impracticable for more than about four equations. Hence alternative methods will be necessary for larger systems.

It is of interest to compare the foregoing solution with that obtained by by Boksenbom and Hood.[2,3] It will be seen that equation 17 takes the place of four equations in their solution. This simplification of solution is a direct result of considering the given system as a particular case of the general multivariable system described previously.

As a simple example of the use of equation 17, consider a given nonsquare system having three inputs and two outputs. Let it be desired to obtain noninteracting controls by the use of a system shown in Fig. 3. Then H and C are

$$H=\begin{vmatrix} h_{11} & h_{12} & h_{13}\\ h_{21} & h_{22} & h_{23}\\ 0 & 0 & 1 \end{vmatrix}$$

$$C=\begin{vmatrix} c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ c_{31} & c_{32} & c_{33} \end{vmatrix}$$

Ratios between the elements of the columns of C are then found by the use of equation 17,

$$\frac{c_{21}}{c_{11}}=\frac{|H_{12}|}{|H_{11}|}=-\frac{h_{21}}{h_{22}}$$

$$\frac{c_{31}}{c_{11}}=\frac{|H_{13}|}{|H_{11}|}=0 \qquad \text{hence } c_{31}=0$$

$$\frac{c_{22}}{c_{12}}=\frac{|H_{22}|}{|H_{21}|}=-\frac{h_{11}}{h_{12}}$$

$$\frac{c_{32}}{c_{12}}=\frac{|H_{23}|}{|H_{21}|}=0 \qquad \text{hence } c_{32}=0$$

$$\frac{c_{13}}{c_{33}}=\frac{|H_{31}|}{|H_{33}|}=\frac{h_{12}h_{23}-h_{13}h_{22}}{h_{11}h_{22}-h_{12}h_{21}}$$

$$\frac{c_{23}}{c_{33}}=\frac{|H_{32}|}{|H_{33}|}=\frac{h_{13}h_{21}-h_{11}h_{23}}{h_{11}h_{22}-h_{12}h_{21}}$$

System Response Equations

The two solutions to the problem of designing the computer for noninteraction are given by equations 11 and 17. They differ in the amount of information used to obtain the solution. Thus the first solution assumes a desired final system response and consequently gives a unique solution for C. On the other hand, the second solution merely assumes the desirability of noninteraction with the result that a nonunique solution is obtained for C. It will now be shown how a knowledge of the desired final system response yields a unique solution for C in the second case.

The final system transfer matrix was given in the previous section of this paper. From the H' matrix it may be seen that the responses of the controlled variables are given by the main diagonal elements of H'. For these elements, $i=j$, and equation 12 becomes

$$h'_{jj}=\sum_{k=1}^{n} h_{jk}c_{kj}-h'_{jj}\sum_{k=1}^{n} h_{jk}c_{kj} \tag{18}$$

This equation relates the diagonal element of the jth column of H' to the elements in the jth column of the C matrix. It was seen previously that, in effect, one element in each column of C was determined. Hence a proper choice of this element will allow any desired h'_{jj} to be obtained. From equation 18, h'_{jj} is given by

$$h'_{jj}=\frac{\sum_{k=1}^{n} h_{jk}c_{kj}}{1+\sum_{k=1}^{n} h_{jk}c_{kj}}$$

or

$$h'_{jj}=\frac{\overline{hc}_{jj}}{1+\overline{hc}_{jj}} \qquad (19)$$

where $\overline{hc}_{jj}$ is the jjth element of the product HC. This relationship enables h'_{jj} to be determined in the case where the computer matrix elements are previously specified.

Alternatively, equation 18 may be written as

$$\sum_{k=1}^{n} h_{jk}c_{kj}=\frac{h'_{jj}}{1-h'_{jj}} \qquad (20)$$

Equation 20, together with equation 17, allows the elements of the jth column of C to be completely determined in the case where the final system response is previously specified.

Disturbance Response Equations

It is now a simple matter to obtain expressions for the disturbance response of the final system. The system of Fig. 4 may be described by the equations

$$Y=H(Z+V) \qquad (21)$$

$$Z=CX+KY \qquad (22)$$

where

Y is a column vector containing as elements the outputs of H

Z is a column vector containing as elements the outputs of C

V is a column vector containing as elements the disturbance inputs, V

X is a column vector containing as elements the inputs to C

H, C, and K have their usual significance

From equations 21 and 22 it follows that

$$Y=HCX+HKY+HV$$

Hence

$$Y=(I-HK)^{-1}HCX+(I-HK)^{-1}HV$$

Thus the disturbance response is given by

$$Y_V=(I-HK)^{-1}HV$$

But

$$H'=(I-HK)^{-1}HC$$

Therefore

$$Y_V=H'C^{-1}V \qquad (23)$$

where, the case of noninteracting controls, H' is a diagonal matrix. Equation 23 enables the component of output signal due to disturbance inputs to be obtained. It should be noted that once noninteraction of controls has been specified and the control responses set at desired values, it is not possible to have any control over the disturbance response of the system.

Conclusions

This paper has illustrated one of the applications of matrix theory in the field of multivariable control systems. The use of previously derived methods enabled an alternative solution to be obtained to the problem of noninteracting controls. Two methods of obtaining the necessary computer elements were derived and it was shown that the system input-output response and disturbance response were readily obtainable.

It is expected that the techniques used in this paper will be of great use in solving other analysis and synthesis problems in the field of linear multivariable control systems.

References

1. INCREMENTAL CONTROL EQUATIONS FOR TANDEM ROLLING MILLS, J. H. Courcoulas, J. M. Ham. *AIEE Transactions*, vol. 75, pt. II, 1956 (Jan. 1957 section), pp. 363–70.

2. GENERAL ALGEBRAIC METHOD APPLIED TO CONTROL ANALYSIS OF COMPLEX ENGINE TYPES, A. S. Boksenbom, R. Hood. *Report 980*, National Advisory Committee for Aeronautics, Washington, D. C., 1950.

3. ENGINEERING CYBERNETICS (book), H. S. Tsien. McGraw-Hill Book Company, Inc., New York, N. Y., 1954.

4. MATRIX FORMULATION OF CONTROL SYSTEM EQUATIONS, A. J. O. Cruickshank. *The Matrix and Tensor Quarterly*, Tensor Club of Great Britain, vol. 5, no. 3, 1955.

5. LINEAR MULTI-VARIABLE CONTROL SYSTEMS, R. J. Kavanagh. *Ph.D. Thesis*, Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada, 1956.

6. THE APPLICATION OF MATRIX METHODS TO MULTI-VARIABLE CONTROL SYSTEMS, R. J. Kavanagh. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 262, Nov. 1956.

◆

Discussion

H. Freeman (Sperry Gyroscope Company, Great Neck, N. Y.): The author is to be congratulated on a very thorough and enlightening presentation of the noninteraction concept in the synthesis of multipole control systems. While the idea of noninteraction was originally developed by Boksenbom and Hood, their work suffered from an overly complicated notation. Mr. Kavanagh has greatly simplified the technique by employing matrix algebra and by taking a more generalized approach. In this respect his work closely parallels work which I recently carried out.[1-3]

Although Mr. Kavanagh has followed closely the approach taken by Boksenbom and Hood, he has unfortunately not touched on some of the questions that were also left unanswered by them. These questions concern the conditions which are necessary to achieve noninteraction (are there any limitations?), the problem of physical realizability and stability considerations.

Thus also the whole question of the purpose or desirability of noninteracting control remains unanswered. Clearly the author does not mean (as he states) that: "Noninteracting controls could therefore enable the speed of the engine to be controlled only by the fuel rate." If the engine is multipole, i.e., multivariable, the speed (one of the outputs) is, by assumption, not exclusively a function of fuel rate. Merely adding external equipment cannot alter this fact. What noninteracting control does imply is that a single system input can be synthesized which will solely and exclusively control speed. This is accomplished by indirectly (according to a calibrated ratio) controlling both fuel rate and propeller angle.

Although it appears that in general it would be desirable to establish one-to-one correspondence between a controlling system input and a controlled output, i.e., noninteracting control, it is not necessary to restrict the synthesis of multipole control systems in this way. Situations can be found where, for reasons of instrumentation, a noninteracting system may be neither desirable nor, in fact, achievable.

A more general approach to the multivariable or multipole system involves merely the requirement that between the outputs and the inputs there exist a stipulated relationship. This relationship can be described as involving magnification, coordinate rotation, and dynamics. Noninteraction is merely a special case of coordinate rotation. While Boksenbom and Hood in their specialized treatment were interested only in noninteraction, a more general treatment should not be so restricted.

Some confusion might arise from the use of the word "system" to denote both the equipment to be controlled and the over-all mechanism which is finally achieved. The word "plant" for the former and "system" for the latter is suggested as an alternative.

REFERENCES

1. MULTIPOLE SAMPLED-DATA CONTROL SYSTEMS, H. Freeman. *Technical Report T-12/B*. Electronics Research Laboratory, Columbia University Engineering Center, New York, N. Y., Sept., 1955.

2. THE SYNTHESIS OF MULTIPOLE CONTROL SYSTEMS, H. Freeman. *Technical Report T-15/B*, *Ibid.*, Apr. 1956.

3. A SYNTHESIS METHOD FOR MULTIPOLE CONTROL SYSTEMS, Herbert Freeman. *AIEE Transactions*, vol. 76, pt. II, Mar. 1957, pp. 28–31.

R. J. Kavanagh: Dr. Freeman has raised certain questions which my paper, being an alternative approach to Boksenbom and Hood's work, did not attempt to answer. However, I have obtained answers to these questions, which are presented elsewhere (see reference 5 of the paper). It certainly is true that there are limitations on the types of systems which may be modified in the foregoing manner in order to achieve noninteracting controls. I have already published a treatment of one such limitation (see reference 6 of the paper). This is the limitation of a suitable row relationship for H. For the problem of noninteraction this limitation reduces to the necessary condition that the m rows of H must be independent (i.e., no row may be formed from a linear combination of some or all of the remaining rows). For a square system it follows that H must be nonsingular. As an example, the system

$$H=\left|\begin{matrix} \frac{1}{1+s} & \frac{2}{1+s} & \frac{1+s}{1+3s} \\ \frac{2s}{1+s} & \frac{1}{1+s} & \frac{1+7s}{1+3s} \\ \frac{1+2s}{1+s} & \frac{3}{1+s} & \frac{2(1+4s)}{1+3s} \end{matrix}\right|$$

cannot be modified by external means in order to obtain noninteracting controls. From a consideration of the number of available degrees of freedom, this result is as might be expected.

I am grateful to Dr. Freeman for his comment concerning the meaning of noninteraction in the example of the turbopropellor engine. It is of course true that we are merely constructing an artificial input quantity for the purposes of the control system.

Noninteracting controls should, of course, be regarded as a particular case of multivariable system synthesis. As such it is apparent that synthesis by Boksenbom and Hood's method (i.e., where $C=-K$) is only one of many synthesis methods. I have devised several other methods of modifying the basic system by linear feedback and premultiplication. These alternative synthesis techniques were derived for use in the general synthesis problem where a given H was to be modified in order to yield a desired H' containing stipulated transfer functions as matrix elements. Analogue computer studies of a noninteracting system synthesized by one of these alternative techniques are at present being made at the University of Toronto.

I do not share Dr. Freeman's apprehension concerning the use of the word "system." It is felt that, provided the qualifiers "basic" and "over-all" are used, no confusion will result. However, Dr. Freeman's comment does emphasize the desirability of a unified nomenclature for this field in view of the increasing interest shown in it. For example, does one refer to multipole control systems or to multivariable control systems?

Vol
Pt.
195

The Case for a Standard System of Railroad Electrification for the American Railroads

T. F. PERKINSON
MEMBER AIEE

FOR MANY YEARS electric-traction engineers have been aware of the advantages that would accrue with system standardization in the heavy electric-traction field, but for various reasons the recognition of a uniform contact system as a standard for mainline electrification is as yet a goal to be attained.

Recent developments in a-c single-phase motive power, and widespread activity in railroad electrification matters outside of the United States, in which activity the use of commercial-frequency a-c power is playing a prominent part, are encouraging domestic railroads and manufacturers to consider the possibility of adopting a standard system for future electrifications in this country. This paper discusses some of the more important aspects of the problem of system standardization for the railroads of the United States.

Why a Standard for Railway Electrification?

The reasons for seeking a single standard system for mainline railroad electrification are not so compelling as those that brought about the standardization of the track gauge. Under the present system of privately owned and independently operated railroads, relatively little interchange or interrunning of motive power between systems takes place and in most of the few instances where circumstances require such interrunning, it is readily accomplished with self-propelled vehicles—either steam or diesel-electric locomotives. In a few cases, interrunning of electric rolling stock does occur and, incidentally, serves to emphasize the advantages of a common, uniform contact system on the one hand, and the disadvantages involved with nonuniform power-collecting systems on the other hand.

Systems Available

Unfortunately, at least from the viewpoint of standardization, there are several different systems available and highly developed by which electric traction can be provided. Each of the systems has its points of superority, and inferiority, and because of and in spite of these characteristics, each is able to perform the task of providing train motive power with a high degree of performance, reliability, and relative economy. Tables I and II indicate the multiplicity and variety of systems operating throughout the world today.

Where local conditions have been governing, the peculiar characteristics of one system in some cases have given this system some advantage over the other, and since there was no impelling reason for modifying the governing local conditions to utilize a standard system when electrification was undertaken, the system most closely fitting the existing conditions was generally adopted.

In some instances where electrification had to be adopted, any one of several systems could have been employed with substantially equally good results, but since there was no commonly recognized standard the decision of system selection was made by committee or commission agreement, by consultants retained for the purpose, or by individual engineers in the employ of the railroad. In each case the agency making the decision advanced technical, economic, or other arguments to demonstrate and prove to management the wisdom of the selection.

Deterrents to Standardization in the United States

Probably the greatest deterrent to the establishment of a standard for railroad electrification in the United States is the understandable absence of general interest on the part of the railroads in the matter; a condition traceable to the development and adoption of the diesel-electric locomotive on a revolutionary and wholesale basis, and to the downward revision of the railroad's position in the over-all transportation industry.

Underlying both of these events is the internal-combustion prime mover: the diesel engine and gas turbine in railroad motive power; and the gasoline, diesel, turbine, and jet engines in the airplane, bus, and private automobile. All of these prime movers now in current use depend on liquid fuels, and so long as such fuels are economically available the self-contained, self-propelled transport vehicle will be the preferred means for moving materials, goods, and people for the large majority of the transport industry.

Self-containment, with resultant mobility and flexibility, accompanied by reasonable economy of operation and facility of acquisition from the financial as well as physical standpoints, are the features that enable the diesel-electric locomotive to hold its dominant position in the railroad motive-power field today.

There are, however, some segments of the railroad transportation activity where self-containment is of less importance than are superior performance and economy of operation. There is little or no valid contradiction to the statement that electrification provides superior performance in railroad motive power, but there are economic reasons for precluding consideration of electrification in many cases. Among these reasons first costs and fixed charges are dominant.

Adoption of a standard system, involving a minimum of fixed-facilities initial costs, which at the same time would permit the use of low-maintenance, moderately priced motive power having developmental potentialities that point to probably lower costs as extended production ensues, would combine to expand the sphere within which electrification would be economically competitive with or even superior to existing forms of self-contained motive power.

With the desirability of system standardization established, the question then arises as to which system, or possibly systems, should be selected as standard.

Early D-C Systems Development

Commercial electric traction in the United States had its inception in the street railways and in this application a high degree of standardization, almost from the beginning, was realized, in that the nominal 600-volt d-c system was

---

Paper 57-16, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 16, 1956; made available for printing November 13, 1956.

T. F. PERKINSON is with the General Electric Company, Schenectady, N. Y.

Table I. System Voltages, Frequencies, Phases for A-C Systems

Voltage	Phases	Frequency	Where Used
Foreign to the United States			
25,000	1	50	France, Turkey, Portugal
20,000	1	50	Germany, France, Belgian Congo
16,000	1	15	Sweden, Norway
16,000	1	$16^{2}/_{3}$	Sweden
16,000	1	25	Sweden
15,000	1	20	Costa Rica
15,000	1	15	Germany
15,000	1	$16^{2}/_{3}$	Austria, Germany, Italy, Switzerland
15,000	1	50	Hungary
11,000	1	$16^{2}/_{3}$	Switzerland
10,000	3	45	Italy
10,000	1	$16^{2}/_{3}$	Norway
10,000	1	25	Sweden
6,600	1	50	England
6,600	1	25	England
6,500	1	25	Austria
6,000	1	50	Germany
6,000	1	25	Germany, Norway
5,500	3	25	Spain
3,700	3	$16^{2}/_{3}$	Italy
3,600	3	$16^{2}/_{3}$	Italy
3,300	1	25	Canada
750	3	40	Switzerland
United States			
11,100	1	25	Pennsylvania Railroad, locomotives and multiple-unit cars; Reading Company, multiple-unit cars; New York, New Haven & Hartford Railroad, locomotives and multiple-unit cars; Virginian Railway, locomotives; Long Island Railroad, locomotives

Table II. System Voltages, D-C Systems

Voltage	Where Used
Foreign to the United States	
3,300	Russia
3,000	Russia, Brazil, Chile, South Africa, Spain, Germany, Belgium, Mexico, Poland, Morocco, Algeria, Italy
2,400	Canada, Chile
1,650	Spain
1,600	Venezuela
1,500	Australia, Brazil, Chile, Czechoslovakia, France, India, Japan, Java, Netherlands, New Zealand, Spain, South Africa, Russia, England, Portugal
1,200	Manchuria, Cuba, Japan, Spain, England
1,000	Germany
840	Switzerland
800	Italy, Germany
750	Switzerland, Italy
660	Chile
650	Cuba, France, Italy, Japan, England
630	England
600	France, Mexico, England
575	Bolivia
United States, Main-Line Railroads Only	
3,000	Milwaukee Railroad, locomotives; Lackawanna Railroad, multiple-unit cars
2,400	Butte, Anaconda & Pacific Railway, locomotives
1,500	Illinois Central Railroad, multiple-unit cars; Chicago South Shore and South Bend Railroad, locomotives and multiple-unit cars
650	New York Central Railroad, locomotives and multiple-unit cars; Long Island Railroad, multiple-unit cars

virtually universally employed. Occasionally an experimental installation was made in which a higher voltage was employed, generally where the urban transit system was integrated with an interurban operation in which the greater distances run made voltages higher than 600 desirable and economical.

As the interurban electric railways developed, the limitations inherent in the relatively low voltage of 600, particularly the disadvantages of excessive voltage drops and their effects on equipment performance, led to the development of higher voltage d-c systems, and the use in some instances of still higher voltage a-c systems.

Interurban electric railways utilizing 1,200 volts of direct current on the contact system came into common usage in the years around 1910, and these were followed in the mid-1910 to 1920 era by short-lived experimental installations carrying 2,400 volts on a third-rail contact system, and another carrying 5,000-volts on an overhead trolley wire.

The 2,400-volt and 3,000-volt systems with overhead-contact systems were developed for heavy main-line railroad traction service in the same period, and the latter appears to be the maximum nominal direct voltage which has been commercially successful and capable of commercial survival.

The "Battle of the Systems"

While the d-c system advocates were reaching toward higher voltages, proponents of the a-c systems were developing these, with the inevitable resultant "battle of the systems" coming into being.

Proponents of the a-c systems emphasized the superior transmission-distribution characteristics afforded by high voltages made practical by the a-c transformer. The penalty for securing these admittedly superior characteristics was exacted by the necessity of using less satisfactory and more costly motive power.

The d-c system advocates stressed the superior characteristics of the series-wound d-c traction motor, which permitted the use of less costly (from the initial and operating standpoints) motive power, with generally superior performance characteristics.

A-C System Development in the United States

The 11,000-volt 25-cycle single-phase system came into being in 1907 with the electrification of the New York, New Haven & Hartford Railroad lines adjacent to the New York terminal area, and with the electrification of a short interurban line of the Erie Railroad between Rochester and Mt. Morris, N. Y. The latter may be considered as the first commercial-frequency electrification of any significance in the United States, since the line was powered through Scott-connected transformers fed from the three-phase 25-cycle commercial power-supply system of the Rochester area.

In only one instance has a greater voltage than the nominal 11,000 volts been utilized on a heavy-traction electrification in the United States; this was on the 22,000-volt 25-cycle single-phase electrification of the Detroit, Toledo and Ironton Railroad, an electrification that was abandoned in 1929.

It is of interest to note that the first main-line railroad electrification in this country, the Baltimore & Ohio's Howard Street Tunnel and Belt Line in Baltimore, was operated at 600 volts, the only electric-traction system that had been commercially developed at the time—1895.

European Developments

So far as European developments are concerned during the early days of electric traction, the experience with d-c systems paralleled closely that in the United States and does so today. The development of the a-c system was, by force of circumstances, somewhat different since Europe had no extensive 25-cycle industrial-power installations and had adopted at an early date a frequency of 50 cycles as a standard commercial-system frequency. In developing a-c electric traction it seemed logical to adopt a frequency, which was an even fraction of the commercial or industrial frequency, for two principal reasons; viz., the provision of frequency ties between the traction system and the commercial power sources was facilitated by the 3-to-1 frequency ratio adopted; and the design, manufacture, and operation of single-phase series-wound traction motors was materially improved with the use of the relatively low frequency of $16^{2}/_{3}$ cps (cycles per second). Simultaneously, contact-system voltages of the order of 15,000 to 16,000 volts were adopted, a limit probably dictated by practical line-to-ground clearance and insulation considerations.

Salient Features of the Two Systems

Both systems, the high-voltage d-c and the higher voltage a-c versions, have their points of superiority.

The d-c series-wound motor is generally acknowledged to be the best available type of motor for traction service, but its commutator sets a practical maximum voltage that can be applied to the motor without incurring undue maintenance costs and sacrifice of reliability in performance. The lack of a practical d-c transformer, or its equivalent, restricts the contact-system voltage to a maximum that can be accepted directly by the traction motors in practical control combinations; this has so far imposed a ceiling of nominally 3,000 volts for d-c systems. This voltage requires substantial conductivity in the contact system, and the conductor requirements with supporting structures are appreciably greater than for the higher voltage a-c systems.

A prime-power supply of alternating current and the use of series motors taking direct current directly from the contact system require wayside conversion equipment which may be motor generators or static rectifiers, both of which involve significant first and operating costs.

However, there are many 3,000-volt and 1,500-volt systems in use and performing successfully in many parts of the world. These systems seem most practical when used on relatively lightly trafficked lines where large amounts of energy in single blocks are not required for normal operations. They are also in successful operation in multiple-unit suburban operations where distances are short but train density is relatively high especially during rush-hour periods. Under very heavy loading conditions involving heavy train operations, either large quantities of contact-system conductors or frequently spaced high-capacity conversion substations would be required. Thus, for the heavy railroading of the type found in many of the American operations, the first costs of distribution and conversion equipment would be materially increased.

Low-frequency a-c systems in general use today require either low-frequency prime-power generation or frequency conversion through relatively costly and power-consuming frequency changers. In addition, special single-phase transmission lines are needed to supply the special wayside transformer substations involved.

While the low-frequency series-wound a-c traction motor has been developed to a degree of high reliability and successful performance characteristics, its initial- and maintenance-cost aspects have always been inferior to those of the d-c motor, and particularly to those of the low-voltage d-c motor. The problems of commutation in the a-c series-wound motor are troublesome, not only in the case of low frequency, but are more so where the motor is required to operate at the higher commercial frequencies of 50 or 60 cps.

The Commercial-Frequency System

The recent developments and improvements in the mercury-arc rectifier and the encouraging prospects that the semiconductor rectifier will be developed in the near future to a point where its use on railroad motive power will be practical, encourage electric-traction engineers to consider a marriage of the two systems, the offspring of which would be the commercial-frequency system employing rectifier motive power with low-voltage d-c traction motors.

Commercial-frequency, or industrial-frequency electrification as it is sometimes termed, has been in the minds of electric-traction engineers for many years—indeed, as noted previously, the Rochester-Mt. Morris line of the Erie Railroad, electrified in 1907, was truly a commercial-frequency system in its era, utilizing the 25-cps three-phase commercial power supply of the area with no special generating or converting equipment except 3-to-2 phase conversion through Scott-connected transformers for 2-directional distribution as single phase on the contact system.

From the viewpoints of efficiency and economy, the ideal commercial-frequency system would permit the use, by the traction motors in motive power, of energy as received from today's established standard commercial power-supply systems as 3-phase 60-cycle power, without change of voltage, phases, frequency, or form, thus eliminating intermediate transformation, and conversion, equipment costs and losses. Practical considerations circumscribe the extent to which this desideratum can be attained.

Voltage transformation between the point of power supply and the traction motor terminals is desirable for personnel safety reasons, control and equipment characteristics, and is relatively easily and efficiently accomplished by the use of suitable transformers at substations and on the motive power.

While 3-phase contact systems can be devised and, indeed, have been and are in use to a limited extent, they are considered obsolete because of the mechanical and electrical complications attendant to the operation and maintenance of a double-contact overhead-conductor system. Thus, any standard a-c system, including the commercial-frequency variety, must perforce operate as a single-phase system as a practical matter.

Phase conversion from three to two, to facilitate distribution to the contact system and to provide some degree of balancing of the prime power-supply system, can be readily accomplished through Scott-connected step-down transformers between the high-tension power supply and the contact system, with little, if any, increase in costs over those involved with the use of normally designed single-phase step-down transformers.

The availability of relatively light-weight a-c/d-c conversion equipment, e.g., the static rectifier, removes the necessity of considering distribution-system frequency from the standpoint of traction motor performance. Hence, there is no valid reason for considering any frequency in the distribution system other than the standard commercial frequency of 60 cps, and no necessity for changing the frequency other than to reduce it to zero in the process of a-c to d-c conversion on the motive power, where rectifier motive power is employed.

Even though direct-motored rolling stock may be employed, no change in frequency is required, providing successful 60-cps a-c motors are available. Some of the European manufacturers have developed and built 50-cycle series-wound single-phase motors for railroad work and, while these are claimed to be successful, the degree of success is questionable particularly in comparison with that obtainable with the d-c series-wound motor. American manufacturers are inclined away from the 60-cycle series-wound single-phase traction motor, particularly since better performance and costs can be secured with the rectifier and d-c motor combination.

Advantages of Commercial Frequency

The principal virtue of the commercial-frequency system is its ability to utilize commercially generated 3-phase 60-cycle energy, without the intermediary of costly, energy-consuming conversion equipment before the traction energy is fed to the motive power. Concomitant with this are the advantages associated in many cases with a relatively high-tension distribution system, which permits the use of a low-cost contact-distribution system.

Question of Unbalance

In some cases, however, a penalty may be exacted for the privilege of eliminating conversion equipment in that problems of prime power-supply unbalance may be introduced, the resolution of which could offset the advantageous elimination of conversion equipment to a material degree. The phase-unbalance problem has been discussed in a recent paper[1] and is not repeated in detail here.

Where commercial-frequency electrifications are in use today, the necessity of special balancing provisions, other than those procurable with contact-system segregation between power-system phases, has not appeared. There is no guarantee that such will be the case in all applications of the commercial-frequency system to American railroading. Traffic patterns, train sizes, types of service, character and capacity of the power-supply system, are some of the factors that will determine whether or not special unbalance-corrective measures are to be taken in specific cases.

Commercial Frequency System as a National Standard

The commercial-frequency system of railroad electrification appears to have sufficient advantages over other systems to justify its recommendation as a national standard for the American railroads. Among serious students of the problem there has never been so close an approach to agreement on one system as there appears to be today. While there are still reservations in the minds of some to the effect that there is no single system that is best for all situations, there appears to be fairly general agreement that the commercial-frequency system could well be recommended as a main-line heavy-traction railroad electrification standard for the United States.

Under the American system of privately owned and operated railroads, recommendation of any standard practice even by such a body as the Association of American Railroads (AAR) does not imply compulsion of adoption by its members, except where matters of equipment (rolling stock) interchange between railroads on a national basis are involved. Consequently, recommendation of the commercial-frequency system as a standard does not mean that, should a railroad decide to electrify, it would have to adopt the recommended standard. Also it does not imply that existing electrifications would be changed over to the recommended standard system, or, if extended, that the recommended system would be installed. It is doubted that a changeover of any of the existing electrifications to the commercial frequency system could be economically justified unless the railroad contemplated extensive new electrification compared with which the existing system would be a relatively small operation.

Possible Exceptions to a Standard System

There still exists some justifiable opinion that, in those situations where a special railroad operation can be isolated from main-line operations involving interchange between railroads, e.g., a suburban commuter service that operates on other than main-line trackage, or that may have to be integrated with the urban transport system of the metropolitan area served by the commuter system, a system other than the commercial-frequency system may be the preferable and more economical method of electrifying. In such instances the d-c systems have much to recommend them.

AAR Active in Promoting Adoption of a Standard

For the past several years, the Electrical Section of the Engineering and Mechanical Divisions of the AAR, assisted by representatives of some of the manufacturers and other industries interested in the subject of railroad electrification, has been studying the broad subject of electrification with a view to determining whether or not a standard, or standards, should be recommended, and if so, what system, or systems, seem most advantageous for adoption to American railroading.

There appears to be little doubt that the commercial-frequency 60-cycle single-phase a-c system is the leading candidate for the position of recommended standard for main-line heavy electric-traction. Whether or not the system can be considered as a single standard to be recommended for all new railroad electrification projects is a matter yet to be resolved.

Reference

1. Power Supply for Commercial-Frequency Railroad Electrification, J. C. Price. *AIEE Transactions*, vol. 75, pt. II, 1956 (Jan. 1957 section), pp. 432–45.

◆

Discussion

L. B. Curtis (Pennsylvania Railroad, Philadelphia, Pa.): Mr. Perkinson's well-written and timely paper is valuable in helping to direct attention to the need in the United States of standardization in railway electrification. As the costs of fuel and diesel-electric locomotive maintenance rise, and train speeds increase, railroad management will have to examine again and more closely the advantages of electrification. Naturally, the lower the first cost of electrification the more attractive it will appear. Standardization means lower costs of the following:

1. Engineering costs. With a standard pattern for new electrifications and major extensions of existing ones, standardization will be possible for structures, catenary systems, substations, locomotives, etc. Basic engineering will have to be performed regarding foundation and pole sizes, spans, catenary make-up, and spacing and details of substations. For similar conditions of load, track curvature, number of tracks, and length of track to be electrified on any railroad, the same design will apply. Therefore, except for specific application and special problems, the engineering costs after the first study will be at an absolute minimum.
2. Material costs. After the initial engineering has determined pole sizes, wire sizes, hardware, apparatus sizes, and other standard parts, these same items can be used on any electrification interchangeably. With all railroads using the same type of material, quantity buying will reduce first costs and also cost and amount of material to be stocked for maintenance. This should stimulate manufacturers to develop better and lower cost materials and apparatus.
3. Construction costs. With a standard specification, standard methods of construction will be developed, labor-saving devices employed, and packaged substations such as the unit type sold, all of which will reduce the cost of installation.
4. Locomotive and multiple-unit car costs. Nothing reduces first cost more than mass production. In the past, electric locomotives have been almost custom made. When standard universal road-type and yard-switcher-type electric locomotives are accepted and used by all railroads, quantity production will greatly reduce both the first cost of motive power and maintenance costs. The same reasoning applies to multiple-unit cars.

As Mr. Perkinson rightfully points out, the 25-kv 60-cycle system offers additional economy as a standard. It would be interesting to know if Mr. Perkinson has any concrete figures he can quote as to the economic advantage of the commercial-frequency system over the present 3,000-volt d-c system, and 11,000-volt 25-cycle system.

E. C. Barnes (Chicago, Milwaukee, St. Paul, and Pacific Railroad, Seattle, Wash.): The author has presented the various reasons for and against the adoption of a standard system of railroad electrification for the United States, and, in doing so, he has not set up a very strong case for standardization. The combination of a commercial-frequency high-trolley voltage, which is reduced and rectified for d-c traction motors, certainly appears to have considerable merit and more so with the developments expected with the semiconductor type of rectifier. However, the problems of phase balance, increased circuit impedance, inductive interference, and the foregoing of the economy of regeneration versus resistance braking, would, in some applications, offset in a marked degree the advantages offered by this system of railroad electrification.

There seems to be a quite prevalent attitude today that all we have to do is decide on a standard system and then many railroads will, as a consequence, electrify. I do not know of any electrification project in this country that is being held back by lack of an established standard. If any system of electrification can be demonstrated to have an over-all economy and performance that is substantially superior to that provided by our present diesel-electric motive power, it is reasonable to assume that many railroads would electrify, but the savings must be substantial and not just nominal, otherwise railroad managements will not approve of the large investment required for any system selected. When such a system of electrification has demonstrated its superiority, it would seem that the adoption of that system as a standard would logically follow.

Tables I and II show the large variation in the character of the power supplied by the contact systems of the different electrifications throughout the world; however, these tables would have been more useful if the miles electrified had been shown for these various electrifications so as to indicate the relative mileage in operation for the different electrification systems.

Modern Computer Analysis for the Design of Steel Mill Control Systems

J. E. REIDER
NONMEMBER AIEE

PHILIP SPERGEL
ASSOCIATE MEMBER AIEE

PROBABLY the most significant development since the war in the measurement and control field is the growth of analog computer techniques for the analysis of control systems for heavy industrial processes. Previously it was necessary to use cut-and-try methods on production equipment to determine the performance of any new system. This procedure was expensive for both the supplier and the customer. The supplier had to develop suitable test models for installation on production equipment without benefit of thorough testing. Field tests extended the supplier's facilities and caused him to do considerable research work away from his laboratory facilities. Equipment failures, component changes or redesigns in basic circuits caused extensive delays and introduced additional co-ordination problems between the supplier and the customer. Also, the customer's costs were not inconsiderable. As a result of his pioneering efforts into new systems, he would encounter increased production problems, generally lower production for the trial periods, and the added nagging worry of having his production facilities serving as the "guinea pig" of new concepts. Even in the field of bacteriology some subjects become fatalities before the serum is perfected. As a result of this risk, prime production or breeding stock is not injected with the newly developed material. Similarly, in industry, in the past, when the customer's best production equipment was eyed by the control "specialists," tests were often permitted on only the limited service equipment and there was still considerable risk in applying a system perfected on older equipment to the newer and faster units.

The advent of the analog computer techniques brought about many desirable changes:

1. The supplier has available, in his own laboratories, complex high-speed tandem mills, wide, fast, Fourdrinier paper machines, the most modern Zee calenders for rubber and plastics, and any number of new machines within a relatively short time after their introduction into industry. He is able to have all of these within very limited floor space, and he has not become a competitor of his end customer in the sense that he does not roll steel, make paper, or calender fabric.
2. Test models can be developed on breadboard-type layouts. These systems can also be thoroughly checked and their performance rated prior to the actual installation of production equipment.
3. The laboratory and production areas are brought together on one floor and the co-ordination problem has been greatly simplified.
4. Component or circuit changes require only minimum effort and costs are consequently minimized.
5. The customer no longer risks his best production facilities on initial developments. He has the assurance that new control systems have been thoroughly tested on equipment capable of duplicating the characteristics of his process.
6. During the initial development stages, he can continue to roll steel, make paper, or calender fabric with a minimum of concern about production.

The chief worry for the supplier is probably in connection with complicated problems of data-gathering.

Coupled with all types of newly developed or war-born electronic developments, analog computer techniques have opened broad new horizons to the suppliers of control systems with the advantages just cited. This has enabled companies to specialize solely as suppliers of control systems. Industrial Nucleonics relies upon analog computer techniques as an important tool in developing measurement and control systems for the various industries. One use of analog computer technique is here discussed in developing controls for the steel industry.

In 1954, the first successful measurement and control system was put into operation on the first stand of a tandem cold mill. At the left in Fig. 1 is shown this original equipment. The measuring head is located between the first and second stands and the controller regulates the first-stand screw positions. The steel industry is now generally familiar with this system.

As a result of the control of strip thickness made possible with this installation, analog computer techniques were used to investigate vernier strip thickness control of the last-stand speed of tandem cold reduction mills.

The analog computer was a convenient tool for simulating this process in the engineering department. It was particularly useful in determining the operation and performance of the control system prior to installation on a mill.

Paper 56-875, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Technical Operations Department for re-presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted August 20, 1956; made available for printing December 11, 1956.

J. E. Reider and Philip Spergel are with Industrial Nucleonics, Columbus, Ohio.

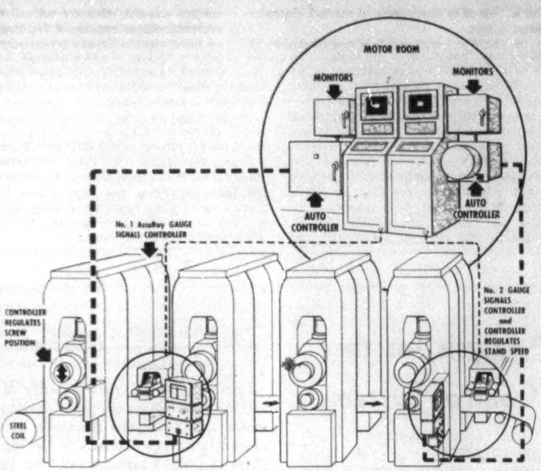

MOTOR ROOM
MONITORS
MONITORS
AUTO CONTROLLER
AUTO CONTROLLER
No. 1 AccuRay GAUGE SIGNALS CONTROLLER
CONTROLLER REGULATES SCREW POSITION
STEEL COIL
No. 2 GAUGE SIGNALS CONTROLLER and CONTROLLER REGULATES STAND SPEED

Fig. 1. **Automatic control of cold reduction mills**

Referring again to Fig. 1, the proposed control system for the regulation of the strip thickness on the output side of the tandem mill may be seen on the last stand. An AccuRay gauge is used to measure the thickness variations in the steel. An error signal generated from this gauge is then fed to a controller which activates a servo motor that drives the vernier rheostat of the fifth-stand speed control system. The decision was made to use this electromechanical addition to the existing control system rather than modify existing circuits with buck or boost components because: 1. the time delay of the servo motor system is negligible in comparison with other lags in this process, 2. this concept made the basic controller circuits adaptable to different types of mill drive control circuits, 3. the cost to the customer is reduced, and 4. the results are comparable with other systems reing modification in the drive circuits.

The vernier rheostat controls, by way of a rotating amplifier, the speed of the fifth-stand drive motor. The speed of the fifth-stand drive motor is varied to counteract the thickness variations of the steel as measured by the AccuRay gauge. The reduction in steel thickness is a function of the difference in speed between the fifth and fourth stands. Fig. 2 is an elementary diagram of the operation of the AccuRay gauge. A radioisotope is used as a source of energy to ionize the gases in the ionization chamber. The steel acts as an absorber of energy and permits more or less energy to enter the chamber, and therefore more or less ionization current is made into a measurable voltage across the high megohm resistor. The voltage across this resistor is then inversely proportional to the thickness of the steel. A high input impedance electronic amplifier is used to convert the small high-impedance signal to usable form for driving the repeat recorder measuring system. A voltage from a repeat slide wire on a recorder is then used as the gauge error signal.

The controller is an electromechanical device. The error signal for the gauge is amplified, the lags in the process are compensated for, and the signal is applied to a servo motor that positions the vernier rheostat.

The dynamic characteristics of the various components of the control system are needed prior to simulation on the computer. The transfer characteristics of the gauge and the process were measured and the necessary calculations were made to determine the dynamic characteristics of these components. "Transfer characteristics" are merely the relationship between the output and input of any circuit or system. For instance, if a control rheostat is moved part of a turn, some increase or decrease in motor speed is obtained within certain time limits. Fig. 3 is a simplified diagram of the field control circuits of the fifth-stand drive motor. The fifth-stand motor field current is a function of the motor field exciter voltage, which in turn is a function of motor field exciter current. Exciter current is varied by varying the resistance in the field of the motor field exciter. There are two rheostats in this field circuit: one is the coarse rheostat which is usually motorized, and the other is the vernier rheostat which is normally under manual control of the fifth-stand operator. It is the vernier rheostat that is presently used to give a fine control of output steel thickness. There are two major delays in this process: The first is the delay of the field current build-up in the main motor fields,

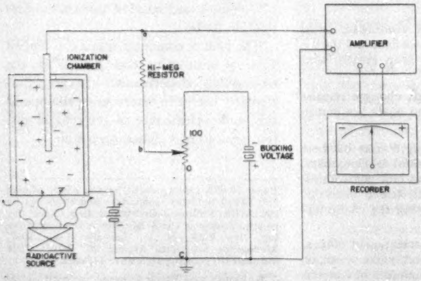

IONIZATION CHAMBER
HI-MEG RESISTOR
AMPLIFIER
a
b
c
100
0
BUCKING VOLTAGE
RECORDER
RADIOACTIVE SOURCE

Fig. 2. **Gauge schematic**

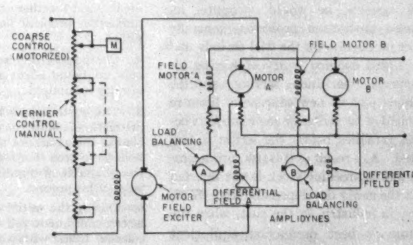

COARSE CONTROL (MOTORIZED)
M
VERNIER CONTROL (MANUAL)
FIELD MOTOR A
MOTOR A
FIELD MOTOR B
MOTOR B
LOAD BALANCING
A
B
DIFFERENTIAL FIELD A
DIFFERENTIAL FIELD B
LOAD BALANCING
MOTOR FIELD EXCITER
AMPLIDYNES

Fig. 3. **Simplified field control**

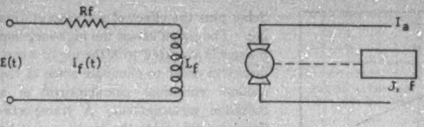
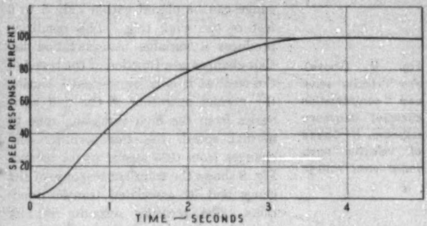


Fig. 4. Elementary circuit



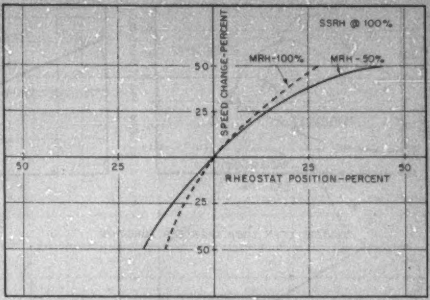


Fig. 5 (left). Speed response to rheostat position step change

Fig. 7 (above). Per-cent speed change versus rheostat position

and the second is the inertia effect of the drive motor armature. Fig. 4 shows an elementary circuit for the calculation of the speed voltage transfer function of a d-c motor controlled by the field voltage. The transfer function shown is a simplification of dynamic functions that would be derived if the delays of the various exciters and the interaction of the load variation were considered.

$$\phi=K_f i_f \tag{1}$$

$$T=K_2\phi \tag{2}$$

$$\frac{\theta(S)}{E(S)}=\frac{\dfrac{K_2}{R_f}\quad\dfrac{K_2}{f}}{(\lambda mS+1)(\lambda fS+1)} \tag{3}$$

$$\lambda m=\frac{J}{F}\quad \lambda f=\frac{L_f}{R_f} \tag{4}$$

where

L_f = field inductance
R_f = field resistance
J = motor inertia
f = motor viscous friction
T = torque
$\theta(S)$ = motor speed transform
i_f = field current
ϕ = field flux
K_2, K_f = constants
$E(S)$ = motor field voltage transform

The following equation shows how complex the output versus input transfer characteristics of the last-stand motor can become if all the delays in the process are accounted for.

$$\Delta\omega(S)=\frac{K_1\qquad\Delta\theta(S)}{(1+T_0S+T_0T_aS^2)\left(1+\dfrac{T_f}{1+A_s}S\right)(1+T_1S)}-\frac{(1+T_aS)K_2\Delta T_L(S)}{(1+T_0S+T_0T_aS^2)} \tag{5}$$

where

$\Delta\theta(S)$ = Laplace transform of the change in vernier rheostat position, degrees
$\Delta T_L(S)$ = Laplace transform of change in motor torque, pound-feet
$\Delta\omega(S)$ = Laplace transform of change in motor speed, rpm
T_0 = mechanical time constant of motor
T_a = armature circuit time constant of motor
T_F = field circuit time constant of motor
A_S = gain factor of motor field exciter differential field
T_1 = field time constant of motor field exciter shunt field
K_1 = rate of change of motor speed with respect to vernier position
K_2 = rate of change of motor speed with respect to motor load

It was decided that an approximation of this complex transfer function would be used for the computer study. To obtain approximate transfer function, a dynamic response of the process was obtained. The transient response of the speed change of the fifth-stand motor as a result of an instantaneous change in position of the vernier rheostat was measured. Fig. 5 is the result of the speed response to the rheostat step position change. The shape of the curve indicates that there are two orders of capacitance-type delays. One is the motor field delay and the other is the motor armature delay. Fig. 6 shows the resultant frequency response as calculated from the data of Fig. 5. The process transfer function of output speed to rheostat position is

$$\frac{\text{output speed}}{\text{rheostat position}}=\frac{K}{(1.33S+1)(0.2S+1)} \tag{6}$$

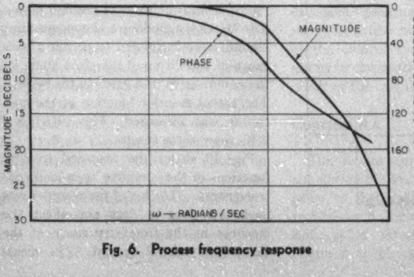


Fig. 6. Process frequency response

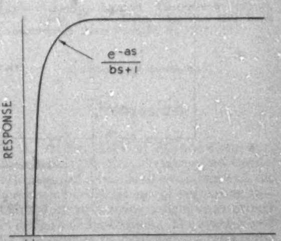


Fig. 8 (right). Gauge transient response

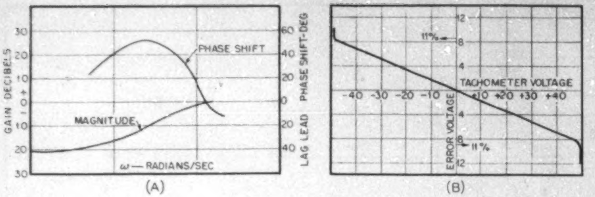


Fig. 9 (above). A—Velocity servo and compensation circuitry: frequency response. B—Speed of velocity servo versus error voltage

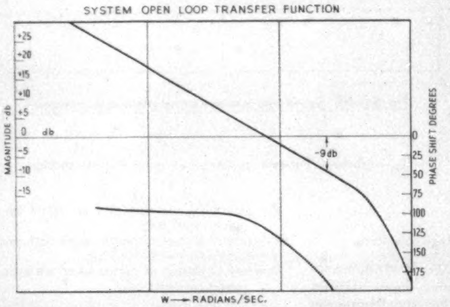


Fig. 10 (left). System open-loop transfer function

The time constants of the delays are 1.33 seconds and 0.2 second. The gain characteristics of the vernier rheostat sensitivity are nonlinear and change as a function of the position of the coarse rheostat of the fifth-stand motor and also vary with the armature voltage of the mill drive motor. Fig. 7 shows how the per-cent speed change sensitivity of the vernier rheostat varies with the setting of the cold reduction mill master pilot and the fifth-stand coarse rheostat. There is an additional nonlinearity in the operation of the vernier rheostat. Due to the construction of the vernier rheostat there is an effective total backlash equivalent to 1/3% of the available speed change.

The AccuRay gauge has one order of delay plus the effect of a transportation lag. The signal across the high-megohm resistor is averaged with the use of a condenser in order to eliminate some of the random variations characteristic of a radiation measurement. A transportation lag exists since the gauge cannot be located along side the work rolls. Because of mechanical considerations, the gauge can be placed within only 5 to 6 feet of the work rolls. This condition produces a variable transportation lag that changes as a function of the speed of the steel as it is being worked. During the normal operation of the mill as it varies from the high threading zone to normal speed, the transportation lag changes from 0.25 second to 0.1 second. Fig. 8 shows the transient response of the gauge and its associated transfer functions. The recorder response is fast compared to the gauge. Its transfer characteristic is

$$\frac{\text{output recorder signal}}{\text{gauge signal}} = \frac{K}{0.0025S^2+0.08S+1} \quad (7)$$

The controller response was determined theoretically from considerations of the various delays in the process and the gauge. Due to the complexity of the existing regulation system of the cold reduction mill, a reset plus rate-action control system was selected. This control system will permit stable operation of the mill for decreasing sensitivity of the regulation system. This type of control system gave minimum interaction with the mill speed regulation loop. There was no need to break into the components of the armature of field circuits of the normal mill speed regulator. To obtain this type of control action, a velocity servo was used to give corrections proportional to the integral of the deviation of sheet thickness from specification. The velocity servo motor shaft was used to drive the vernier rheostat. Compensation circuits were calculated and designed to compensate for the field and motor armature delays of the mill motor. The velocity servo loop and compensating circuits were calculated to permit a wide range of stability and therefore allow for the nonlinearity that exists in the system. The actual transfer function of the controller was measured. Figs. 9(A) and 9(B) show these results.

Fig. 10 shows the cascaded transfer functions of the complete open-loop control system. The lag of the recorder was neglected since it does not affect the response in the frequency range of the closed-loop control system. The meas-

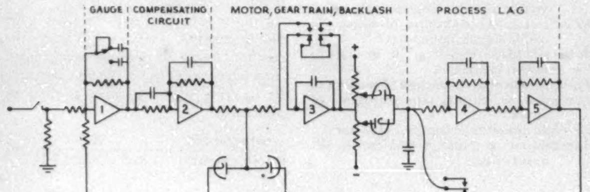
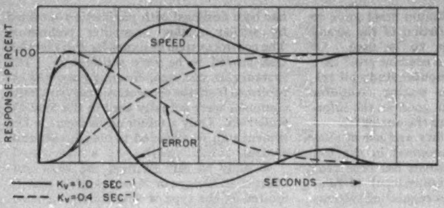
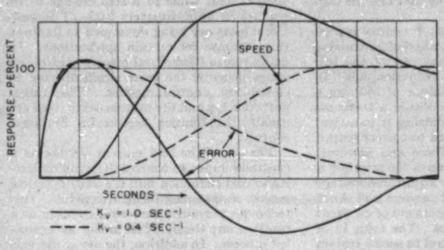


Fig. 11. Control system computer setup





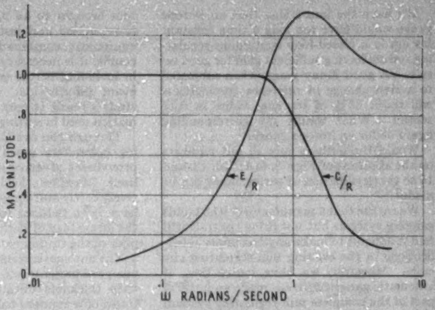


Fig. 12 (above, left). Transient response of control system with compensation

Fig. 13 (left). Transient response of control system without compensation

Fig. 14 (above). Closed-loop frequency response

ured transfer function of the process was used. The effect of the nonlinearity of the vernier rheostat will be to shift the magnitude curve of the frequency response up or down, depending on the velocity constant of the system as determined by the vernier rheostat sensitivity. The velocity constant or control system gain was calculated to ensure complete system stability.

Knowledge of the various delays of the control system components permitted the simulation of the complete process on an analog computer. The purpose of this simulation was then to check out the calculated performance and stability of the system.

Fig. 11 is a block diagram of the computer setup used for this simulation. The computer was set up in the form of transfer function blocks so that the performance of each component of the system could be observed independently. The nonlinearities incorporated in the simulation were those of control motor speed saturation and backlash of the vernier rheostat. Various transient responses to disturbances in the process were obtained. The purpose of these measurements was to determine the effect of the various parameters in the process. Fig. 12 shows the control system response to an abrupt 1.5% thickness change in the process as a function of the system gain. The vernier sensitivity, and therefore the system gain, will vary as a function of the speed of the mill as seen from Fig. 7. The transient response as shown in Fig. 12 indicates that the system has a sufficient margin of stability and will not produce undesirable oscillations during control operation. Fig. 13 shows the transient response of the system as a function of gain change and no compensating circuits in the controller. It is observed from these results that compensation is needed if a high-performance control system is desired. With the compensation circuits, the error due to the transient will be reduced in one third the time it takes the control system that did not have compensation. Since the control system was designed with a large margin of stability, the effect of varying the transportation lag from 0.10 to 0.25 second does not change the transient response of the compensated system. The small backlash of ±1/6% of process does not affect system performance.

The frequency response of the control system, as simulated on the computer, was measured. The results indicate no excessive "peaking" and give a figure of merit as to the performance capabilities of the control system to random variations in the process. Fig. 14 shows the result of the frequency response.

Conclusions

The analog computer permitted the simulation of the complete control system prior to installation of an actual controller on a high-speed cold reduction mill. This simulation permitted the engineer to check the calculations and performance of the closed-loop control system. In this way an additional safety factor was obtained, and thus a safe, reliable performance of the controller when it was installed in the steel mill was ensured.

This is one example of how computer techniques can be utilized and it shows what can be done by a supplier prior to installing control systems on production equipment. These techniques also make possible customizing of basic control systems in order to give maximum results from any existing equipment.

References

1. PRINCIPLES OF SERVO MECHANISMS (book), G. S. Brown, D. P. Campbell. John Wiley & Sons, Inc., New York, N. Y., 1948.

2. TRANSIENTS IN LINEAR SYSTEMS (book), M. F. Gardner, J. L. Barnes. John Wiley & Sons, Inc., 1942.

3. SERVOMECHANISMS AND REGULATING SYSTEM DESIGN, VOL. 2 (book), H. Chestnut, R. W. Mayer. John Wiley & Sons, Inc., 1955.

4. ELECTRONIC ANALOG COMPUTERS (book), G. A. Korn, T. M. Korn. McGraw-Hill Book Company, Inc., New York, N. Y., 1952.

◆

Discussion

R. A. Phillips (General Electric Company, Schenectady, N. Y.): We agree that the development of the analog computer and the concurrent development of regulating system theory have made a significant change in design and application. The system used by the authors to illustrate the techniques is of considerable interest to us.

We have the impression that an isotope gauge would have too long a time constant for use in a closed-loop continuous-regulating system having sufficient gain for good results. A good X-ray gauge when subjected to a step change in reference to thickness will reach 63% of its final value in 0.15 second. What would the corresponding figure be for an isotope gauge?

Would the authors care to put numbers on the abscissas of Figs. 8, 9(A), 10, 12, and 13 so that evaluation of performance can be made?

We realize that a manufacturer who builds gauging systems but not drive controls will find it difficult to make any but minor modifications in the existing mill-regulating circuits. However, we have found that, if automatic gauge control is made an integral part of the complete mill-regulating system, the resulting improved gauge tolerance justifies the effort.

On an existing installation similar to the authors', where control of motor speed is by field, we studied the problem on an analog computer and found that changes in the existing regulating system were desirable. The principal change was to replace the field current vernier with a field current regulator. Our study showed that the added forcing and the effective decrease in the motor field time constant permitted system gains sufficient to decrease response time to about one half that shown in Fig. 12 (assuming that divisions are one second in the figure). Additional improvement in performance may be obtained by modification of the winding reel control.

E. H. Browning (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This paper is a contribution to the literature. Once a system has been studied and represented in analog form, the effects of adjustments can be readily illustrated and the system brought to an optimum state prior to construction and installation of the actual equipment combination to be used. Of course, it is necessary to exercise great care in including in the computer study all relevant information; an analog computer study's result is only as good as the information used in setting up the analog.

Through the availability and use of analog computers much progress in the improvement of drive systems has been realized. Further, field adjustment of relatively costly and complex regulating systems have been reduced to a minimum state to the advantage of both the user and the supplier of the equipment.

The automatic strip gauge control system analyzed presents one method of minimizing strip thickness deviations out of the last stand of a tandem cold reduction mill. In its basic form, the technique of utilizing a radiation-type gauge to obtain a thickness deviation signal, and utilizing it to adjust the last-stand drive speed to correct output thickness, is sound. There are, however, many differences in the ways the principle is applied in detail. The adjustable voltage systems used to control tandem mill stand drive motors and generators are of excellent regulating characteristics. The tying in of an automatic gauge control to such a system must be carefully done and a high degree of co-ordination is necessary, to result in an integrated system of expected quality.

J. E. Reider and Philip Spergel: We thank Mr. Browning and Mr. Phillips for their discussions. The intent of this paper was to present a technique of analyzing a tandem cold reduction mill control system for thickness control on an analog computer. The various components of the control system were described and the analog of these was set up on the computer.

A thickness control system installation has been designed with parameters obtained by similar analog computer techniques The control system designed for this installation proved to be very satisfactory. The system gain and compensating networks established from the theoretical analysis on the computer were actually used in the final installation. Only a minor adjustment in the system gain was needed to obtain optimum control.

In reply to Mr. Phillips, the time constant of an isotope or X-ray gauge is mainly limited by the type of detector that is used. Detectors used in isotope gauges permit 63% of final value to a step change to be reached in approximately 0.05–0.1 second. Techniques are being developed to improve this response for certain applications. In many cases of industrial process control the gauge delay is the least significant lag in closed-loop control system. The transportation lag and the process delay itself are usually the limiting factors for optimum control.

The use of the field current vernier as a controller input to control the speed of last-stand cold reduction mill is a very practical means of obtaining integral control. This technique permits a zero error to occur as a result of any steady offset in the uncontrolled process. In addition, the use of the vernier permitted a wide range of variation of steel thickness to be corrected without the need to readjust the coarse rheostat. The design techniques used permitted the steel mill operator to utilize his well-established rolling practice.

The paper was intended to present a technique of control analysis. No direct comparison of the results of this computer study with those obtained by Mr. Phillips' should be made, since each mill installation should be studied separately. The system transient response will vary and it depends upon the particular characteristics of the mill under consideration.

An Ignitron Contactor for Direct-Press Drives

D. L. PETTIT
MEMBER AIEE

R. MONTROSS
ASSOCIATE MEMBER AIEE

THE USE of direct drives for large draw presses in industry was originally brought about by the continued demand for higher rates of production as well as higher capacities in the presses themselves. The usual method of allowing the main drive motor to run continuously while the press is started and stopped by the operator controlling a clutch has become extremely costly because of production interruptions for clutch maintenance.

By gearing the drive motor directly to the crank of the press, and starting and stopping the motor for each operating cycle, the problem of clutch maintenance is eliminated. However, as motor horsepowers increase it becomes evident that the problem of maintenance has not been eliminated, but merely transferred from the clutch to the motor and its controller.

Since the motor is started and stopped for each stroke of the press, a large contactor must close and open the circuit for each cycle of operation. The high rate of operation causes rapid erosion of the contact tips, and contributes to relatively short mechanical life of the contactor. Shock loading of the gear train each time the motor is connected to the line causes damage to gears, shafts, and keyways, eventually requiring complete overhaul of the press.

Additional contributions to contactor distress and motor maintenance are produced during die setup procedures, when the operator may repeatedly inch the press through most of a complete press cycle. Under these circumstances arcing is present at the contactor tips. The motor may be subjected to very severe transients which produce additional thermal and mechanical stresses within rotor

---

Paper 57-191, recommended by the AIEE Industrial Control Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 23, 1956; made available for printing December 17, 1956.

D. L. PETTIT and R. MONTROSS are with the Square D Company, Milwaukee, Wis.

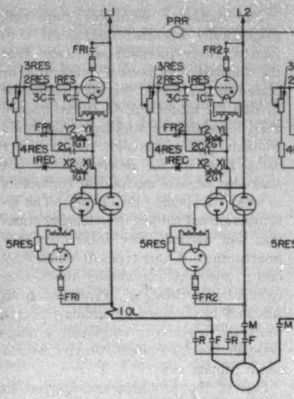


Fig. 1. Simplified diagram of controller

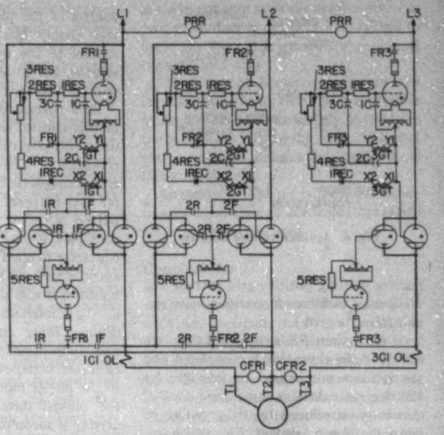


Fig. 3. Reversing using ignitron

bars, rings, and phase windings. Inrush currents to the motor may be sufficient to produce welding of the main contactor, which shuts down the press until maintenance crews can adjust or replace the contacts.

The ignitron contactor for direct press drives is a new approach to relieving these difficulties. By connecting a pair of ignitrons back to back in each of the three phases, a contactor is obtained which is capable of closing the circuit for each operating cycle and still render long, trouble-free service.

Use of ignitrons as contactors permits utilization of their useful characteristic of reducing output voltage by retarding their firing angle, commonly known as phase back. By incorporating phase-back action with slope control, the motor terminal voltages is permitted to build up to nominal value in a few cycles. Thus a cushioned start of the motor is obtained with negligible increase in starting time.

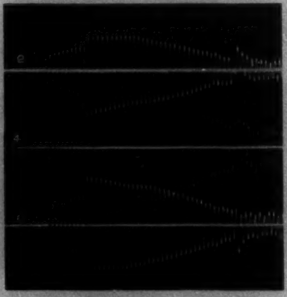


Fig. 2. Oscillograph trace of accelerating current

The cushioning action of the phase-back circuit has several useful functions. It minimizes the effect of the d-c transient in the motor, thereby reducing the component of braking torque caused by this transient. Slack in the gear train is taken up softly instead of with a sharp hammer blow as would occur when the motor is started at full line voltage. Magnetic forces within the motor are reduced at the moment of voltage application, lessening stress to the rotor conductors and reducing working of the phase windings to the point where motor life is multiplied. Smaller increments of movement are possible when inching the press because the small relays used to control the ignition circuits are much faster in response then the large mechanical contactors normally used to control the motor current directly. This makes the press respond more quickly for inching operations.

Controller Operation

A simplified diagram of the controller is shown in Fig. 1. Nonessential features such as transformer primaries, safety circuits, etc., are omitted. It will be noticed that each phase employs two ignitrons connected back to back. Associated with each ignitron is a thyratron which is used to control the firing of the ignitor. Operation is obtained by closing the contacts of firing relay *FR* which is controlled by the operator.

Each of the thyratrons in the lower part of Fig. 1 controls an ignitor connected through relay contacts *FR*. The thyratrons function as diodes. Ignition is obtained by closing the contacts *FR*, provided there is a path for the current to flow through one or more of the thyratrons in the upper part of Fig. 1.

The thyratrons in the upper part of Fig. 1 control all the currents flowing to the motor, while those below control the current from the motor. Since any current flowing at all must simultaneously flow both to and from the motor, at least two thyratrons or ignitrons must be conducting simultaneously.

By the same token, control of the grids of the upper three thyratrons controls the current to the motor, and the remaining three thyratrons and ignitrons, which conduct the current from the motor, are slaves and can only follow. The manner in which the upper three thyratrons control the current to the motor is as follows.

Referring to Fig. 1, consider the condition with relay *FR* de-energized. Capacitor *3C* is by-passed through *2RES*, which is of low ohmic value and limits the rate of discharging *3C* later in the sequence. The grid, therefore, is at the potential generated by the *Y1–Y2* winding of transformer *1GT* (*2GT* and *3GT*). The transformer primaries are connected to the supply line to give grid voltages which lag respective plate voltages by 120 degrees. (*1RES* is the standard grid

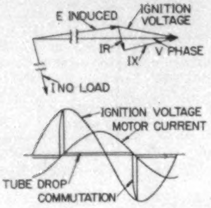


Fig. 4. Ignition voltage diagram

resistor limiting positive grid current and *1C* is used to filter off spurious noise signals from the grid.)

At the instant *FR* is energized the plate circuit of the thyratron is completed and the ignitors are caused to fire late by 120 degrees. At this same time an *R-C* circuit is introduced by the opening of normally closed contact *FR*, which removes the discharge circuit from *3C*. *3C* then takes on a positive charge from transformer winding *X1–X2* through rectifier *1REC* and resistors *4RES* and *3RES*. The latter is adjustable to control the rate of charging.

As capacitor *3C* takes on positive charge, the grid voltage, now made up of the sum of this charge and transformer voltage *Y1–Y2* increases in the positive direction. As it does so, the thyratrons and the ignitors are fired earlier in the cycle until positive capacitor voltage finally exceds peak transformer *Y1–Y2* negative voltage. The entire system is then firing full on. This sequence takes place in a few cycles and is referred to as slope control.

It is interesting to note that the control by phase-back network of three of the ignitrons, while the remaining three are phased full on, gives full control of the alternating current delivered to the motor. One way of explaining this might be that the currents flowing to the motor must equal the currents flowing away from the motor. Since the three upper tubes control the currents flowing to the motor, the remaining three tubes are slaves and must follow.

In Fig. 2 the oscillograph trace shows the initial current, the slope for a few cycles, and then the application of full current to accelerate the motor to full speed. It will be noted that there is a step change in the current values as the full-current condition is reached. This occurs because the combination of a d-c bias and an a-c grid control voltage gives a range of only 90 degrees. As the d-c bias, in the positive direction, approaches the peak value of the negative a-c wave, control of the phase back is lost. This step change has not been found detrimental to the operation of the control scheme, but if desired can be eliminated with an alternate firing arrangement.

In a normal press-control application of this nature, reversing of the motor is used only for die setup and emergency conditions. Hence a mechanically operated magnetic contactor is satisfactory for this purpose. Since it must be employed in series with the main ignitron contactor, it may also be used as a safety contactor to open the motor circuit in the event of ignitron failure. Reversing with ignitrons is also possible as shown in Fig. 3. It is interesting to note that no additional firing circuits are required for this connection. The six existing circuits are switched with small contactors to the proper pairs of ignitrons to obtain the desired direction of rotation. Reversing is accomplished by interchange of only lines *L1* and *L2*. This eliminates any transfer switching requirements for the circuitry of *L3* so that a single pair of ignitrons will suffice for control of this line. Observe that reversing with ignitrons is obtained with the basic additions of two pairs of ignitrons and the small contactors necessary for switching control to the proper ignitron pairs.

For contactor sequences employing magnetically operated mechanical contactors, safety precautions demand that the forward and reverse contactors be placed in series with the main contactor so that the motor circuit can be opened even if welding or sticking of the main contacts is encountered. A controller using ignitrons for reversing does not normally require mechanical contactors in series for opening the circuit. The ignitrons are not subject to welding or sticking, and normally are inherently self-interrupting. Other types of failure may occur, such as continuous conduction caused by overheating. Protection from this may be obtained by adding an extra contactor in series or employing a magnetically operated circuit breaker with a shunt tripping coil.

One of the problems encountered in ignitron control of motors is to maintain operation during possible no-load or regenerative conditions that may be found in the press cycle. The motor ignitrons may be called upon to conduct when only the no-load exciting current of the motor is being drawn. Fortunately, some relief to this problem is inherent in the design of the large-horsepower motors used, which generally are highly saturated, causing their no-load exciting currents to

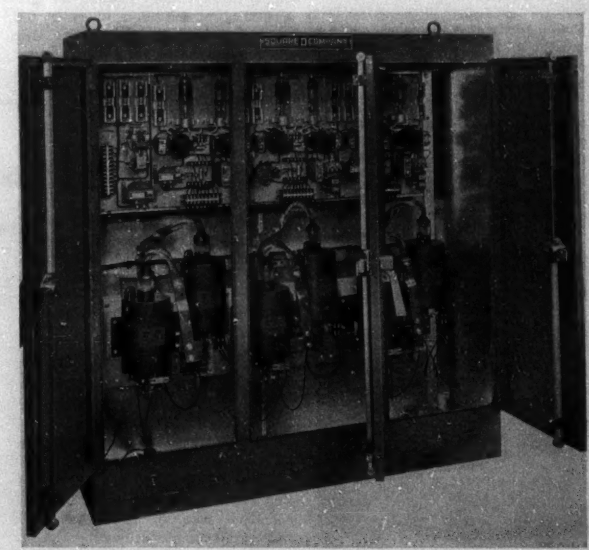


Fig. 5. Main contactor panel with firing circuits

be a large percentage of their full load currents. This high no-load current provides sufficient magnitude to guarantee firing of the ignitors even during no-load or regenerative conditions. Firing circuit variations have been evolved which use thyratron grid hold-off to prevent conduction through the ignitors when motor currents are obtained below the satisfactory level for firing of the ignitrons. These circuits have not been required to date because the motor sizes so far encountered have not reached this lower limit.

While minimum ignitor current is important, a certain voltage is also required to insure ignition. This voltage available to the ignitor circuit is a function of the counter voltage developed by the motor. During starting and acceleration, there is no problem because the counter voltage is reduced in magnitude by the large impedance drops in the motor. Fortunately, direct-drive press motors are usually applied at 440 volts so that sufficient difference voltage still exists between the line and motor counter voltage during running conditions to permit this type of control to operate.

When examining the firing voltage in back-to-back operation, the angle between this voltage and the current becomes important. Reference to Fig. 4 will show that this angle is essentially governed by the impedance angle of the motor. For a given motor this angle remains relatively constant. The voltage available to fire is the difference between the line voltage and the counter voltage of the motor on a vector basis for each phase. In back-to-back operation only one tube in the pair can conduct at a time. As one tube ceases to fire, the impedance drop voltage appears almost instantly to fire the ignitor in the other tube in the pair. This commutation takes place as this voltage has progressed well up towards the maximum value on its sinusoidal curve. It is this set of conditions which makes the load-fired ignitor feasible for this application. Motors normally applied for direct press drives have been found to have impedance angles and impedance drops of proper magnitudes for success.

Several precautions must be observed in the application of this controller. Phase-reversal relays are required, as shown in Figs. 1 and 3, to guarantee proper phase sequence on the input terminals so that the phase-back circuits are properly oriented. Proper connection to insure correct direction of motor rotation must be done at the motor terminals. Relays *CFR1* and *CFR2* (Fig. 3) across the motor terminals are so connected as to trip the circuit breaker if the tubes continue to fire after the firing relay is de-energized. Such an event would maintain one or both of these relays energized, and a timer, in turn, would trip the breaker. Overload relays of the magnetic type are employed in the motor circuit to give an additional measure of protection in the event that large d-c components appear in the motor circuit due to failure to fire of any one of the ignitrons over several consecutive cycles.

Where motor horsepower is too low for the ignitor firing circuits, it is possible to employ thyratrons directly in the lines to the motor featuring the same method of control except without the problem of firing the ignitors.

An incidental benefit is the great reduction in noise level of controller operation when ignitrons are used.

Conclusions

An ignitron contactor panel complete with firing circuits is shown in Fig. 5. The cushioning action of the phase-back and slope-control features reduces the impact of shock loading to the motor, gears, keyways, and other mechanical parts of the drive, in turn reducing maintenance requirements. The response characteristic is greatly improved over that obtained with electrically operated mechanical contactors.

Because of the absence of large moving parts, and the use of static components, this controller requires little maintenance and offers long, trouble-free service. Drastic reductions in down time and attending production losses are obtained.

These important improvements are obtained at the relatively minor additional cost of the initial installation and of providing water for cooling the ignitrons.

◆

Discussion

Joe S. Kirk (National Electronics, Inc., Geneva, Ill.): I believe the authors have thoroughly covered the advantages and operation of the ignitron contactor for press drives and several of the problems involved in its design. The selection of the proper ignitrons presented additional problems.

First it was necessary to determine the size of the ignitrons required. An oscillogram of the motor current was obtained for one cycle of the press. From this oscillogram and the per-cent duty it was possible to calculate the load and establish the size of the ignitron needed.

Second it was necessary to develop an ignitor suitable for the application. Previous attempts to use ignitrons have been unsuccessful because the ignitors were not made for this type of operation. There are two types of ignitors being currently used: those for resistance welding ignitrons, and those for rectifier ignitrons.

The order of magnitude of the impedance voltage available for firing the ignitor mentioned in the paper indicated that the welding ignitor should be used. But, the per-cent duty was much higher than normally encountered in welding applications. It was felt that the shortened life obtained at this high duty was sufficient reason for not using this type of ignitor.

The ignitor used in rectifier ignitrons is designed for long life at high per-cent duty, but the voltage required for ignition of most rectifier ignitors was sufficiently high to prevent their use. For the first ignitrons used in this application, rectifier ignitors were specially selected for low ignition voltage. As the ignitor ignition voltage is of a statistical nature, it was expected that there would be a few misfires due to the low ignition voltage available. When the tubes were installed in the equipment, it was found that several misfires occurred during each cycle of the press. It was felt that a better ignitor was required even though the misfires apparently did not interfere with the press operation to any great extent.

An improved ignitor was developed which provided low ignition voltage and long life. Tests taken during operation show that misfires seldom occur. An installation has been in service for over a year without any signs of tube trouble. On the basis of resistance measurements, the ignitors show no sign of wetting or of excessive erosion after a year of service.

Protecting A-C Motors with Low-Voltage Air-Circuit-Breaker Series Trips

F. P. BRIGHTMAN
MEMBER AIEE

R. R. McGEE
MEMBER AIEE

P. J. REIFSCHNEIDER
MEMBER AIEE

LOW-VOLTAGE air circuit breakers are sometimes used alone as motor starters, since they provide in a single device the various elements of a motor-branch circuit as required by the National Electrical Code. The purpose of this paper is to discuss the effectiveness of the direct-acting series trips used on such breakers as a means of protecting squirrel-cage motors in the event of: 1. Locked or stalled rotor, either during starting or running. 2. Overheating due to prolonged moderate overloads. 3. Short circuits in the motor or in connections to it.

The possibilities of making firm recommendations as to the suitability of low-voltage circuit-breaker direct-acting series trips for motor protection is complicated by: 1. Wide variations in the time-current heating characteristics of the motors, particularly during the starting and stalled-rotor condition. 2. The great variety of time-current tripping characteristics that can be obtained with the present-day circuit-breaker mechanisms, many of which have adjustable time-delay and current pickup.

Consequently, the authors have attempted to present only a broad, general picture of circuit-breaker and motor time-current characteristics, including graphical demonstration as to how well the circuit-breaker tripping mechanism's characteristics conform to the motor-heating requirements. It is hoped that this general discussion of the various factors involved will: 1. Arouse sufficient interest on the part of suppliers and/or users to persuade them to check the suitability of a proposed installation. 2. Provide sufficient background information to enable the supplier and/or user to determine whether or not the protection afforded by a direct-acting series trip is acceptable for a specific application.

For obvious reasons, the authors have used the time-current characteristics of General Electric Company equipment with which they are most familiar as a basis for the discussion. It is believed, however, that the curves and other data presented are typical of motors and circuit breakers made by others.

Motor Heating Characteristics

Broadly speaking, there are two periods in the operation of motors when overheating protection is needed: normal-running-overload condition, and stalled-rotor condition.

NORMAL-RUNNING-OVERLOAD CONDITION

Under the heading of running-overload protection, the motors may be divided into three classes having definite American Standards Association and NEMA (National Electrical Manufacturers Association) standards for temperature rise, namely: 1. 50 C (degrees centigrade) continuous with no service factor, 2. 40 C continuous with 115% service factor, and 3. 55 C continuous with no service factor (enclosed motors).

According to NEMA standards, these temperature rises are based on an ambient of 40 C. These recommended temperature limits have been adopted as industry standards for class *A* insulation on the grounds that insulation operating at the total temperature permitted would have a long-life expectancy. In this connection, the question of the probability of the occurrence of moderate overloads is an important factor in evaluating the effectiveness of protection afforded by the series-trip devices. Certain types of loads such as centrifugal pumps and fans are not likely to be overloaded, assuming that the motor was properly selected for the application. On the other hand, reciprocating pumps, conveyors, machine tools, coal pulverizers, crushers, etc., can impose overloads of moderate magnitude which may not be easily detectable by the operator.

STALLED-ROTOR CONDITION

A motor rotor may stall either during starting or running due to extreme overload, low voltage, or freezing of the bearings, etc. When stalled, the motor draws symmetrical locked-rotor (initial-starting) current for the duration of the stalled condition. Since stalled rotor is an abnormal condition, which is not expected very often, the rotor temperature is allowed to go considerably higher than for normal running conditions. There are no accepted standards for the maximum permissible temperature rise permitted on the motor stator or rotor (whichever is the limiting item).

The speed of the machine and the type of rotor construction, i.e., whether it is cast aluminum or copper bars with brazed end rings, are the principal factors affecting the permissible stalled-rotor time. These factors and others are taken into account by the engineer when he designs the motor and shows in the form of a heating curve that the motor can safely withstand a certain number of seconds with a stalled rotor.

The present General Electric general-purpose, 10- to 500-hp (horsepower) squirrel-cage induction motors are designed to allow approximately 20 seconds at stalled rotor with six times normal current flowing. In special cases, however, the stalled-rotor time will be considerably longer than that, while in others it may be substantially shorter if the motor has to be designed to meet certain requirements necessitating a sacrifice somewhere.

Relationship of Allowable Starting and Stalled-Rotor Times

Stator-winding temperature under starting and stalled-rotor conditions generally is the limiting factor in the design of small motors (about 75 hp and below), whereas rotor heating is the limiting factor in the larger motors. Since the starting current of normal-slip, squirrel-cage motors does not drop substantially until the motor is nearly up to speed, as shown in Fig. 1, the stator winding is subjected to essentially the same current magnitude during starting as it would be when stalled. Hence, a small motor whose stator is designed with sufficient capacity to handle a slow-starting, high-inertia load can stand stalled-rotor condition for approximately the same period, neglecting the effect of ventilation when the motor is starting.

When the rotor is the limiting factor, however, there may be a considerable difference between the time that a motor can safely withstand starting current compared with the time it can withstand essentially the same magnitude of current when stalled. The reasons for this are as given in the following:

Paper 56-652, recommended by the AIEE Industrial Control Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, San Francisco, Calif., June 25–29, 1956, and re-presented at the AIEE Fall General Meeting, Chicago, Ill., October 1–5, 1956. Manuscript submitted October 21, 1955; made available for printing July 24, 1956.

F. P. BRIGHTMAN, R. R. MCGEE and P. J. REIFSCHNEIDER are all with the General Electric Company, in Schenectady, N. Y., Plainville, Conn., and Philadelphia, Pa., respectively.

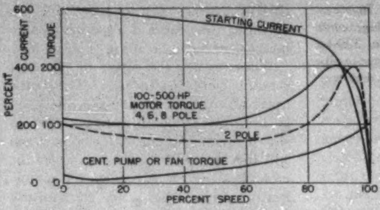
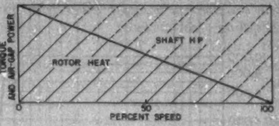


Fig. 1 (left). Squirrel-cage induction motor, torque and starting current characteristics

Fig. 2 (right). Division of input power to squirrel-cage induction-motor rotor



1. The rotor resistance decreases substantially, as a normal-torque squirrel-cage motor comes up to speed. The change in rotor resistance R_2 may be as much as 4 to 1, and the heating varies proportionately, since it is determined by I^2R_2.

2. There is a small reduction in starting current.

3. Windage has little effect on rotor heating when the motor starts quickly, but there is more time for the heat to be carried away as the starting time increases.

4. The changing division of air-gap power, rotor heat, and shaft horsepower output, as shown in Fig. 2.

Fig. 1 shows typical motor torque curves for 100- to 500-hp, 2-, 4-, 6-, and 8-pole motors in comparison with a typical load-torque curve for a centrifugal pump or fan drive. This figure also shows a typical starting-current pattern.

A motor with an easy starting load and ample margin between the available motor torque (4-, 6-, 8-pole motor curve), and the torque required by the load (pump or fan-load curve), Fig. 1, comes up to speed in a very few seconds. In such cases, an allowable stalled-rotor time of 20 seconds exceeds the required starting time, and the protection problem is simple.

If the motor has substantially lower torque with respect to the required torque, e.g., the 2-pole motor curve in Fig. 1, the net torque available for acceleration is reduced, and the motor requires a longer time to bring the load up to speed. If the required starting time for a given installation is long enough, the starting time exceeds the allowable stalled-rotor time, as explained previously. Such a condition is indicated by the motor-heating and circuit-breaker time-current curves *1A* in Fig. 3. This characteristic has an important bearing on the degree of protection a given device can afford to a motor, as explained in the next section.

Motor Time-Current Curves

The motor-heating curve in Fig. 4 shows typical time-current performance for these general-purpose motors when designed for 115% service factor. In reality, this motor-heating curve should be broken into two parts: the upper or low-current end showing the temperature rises at moderate sustained overloads (normal running condition), and the high-current end showing the allowable time before a dangerous temperature is reached with a stalled rotor. Incidentally, it should be borne in mind that the increased temperatures due to moderate overloads do not mean immediate destruction of the motor or even serious damage, but rather that the insulation life may be shortened if these temperatures are allowed to persist indefinitely.

Motor-Overheating Protection Afforded by Direct-Acting Series Trips on Large Low-Voltage Air Circuit Breakers

Fig. 3 shows the typical motor-heating curve for 10- to 500-hp normal-torque squirrel-cage motors in combination with three possible time-current curves obtainable with the new *EC-2* series tripping mechanism, Fig. 5, for General Electric *AK-1-15*, *AK-1-25*, and *AK-1-50* circuit breakers; see Fig. 6.

The current pickup settings of the long-

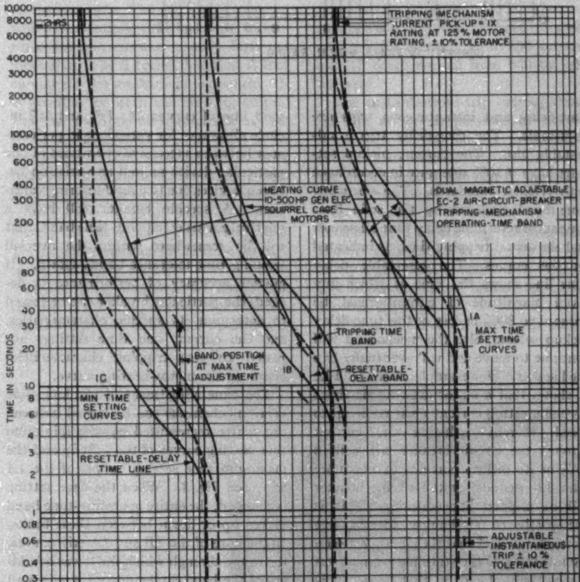


Fig. 3. Time-current curves of induction motors and EC-2 tripping mechanism

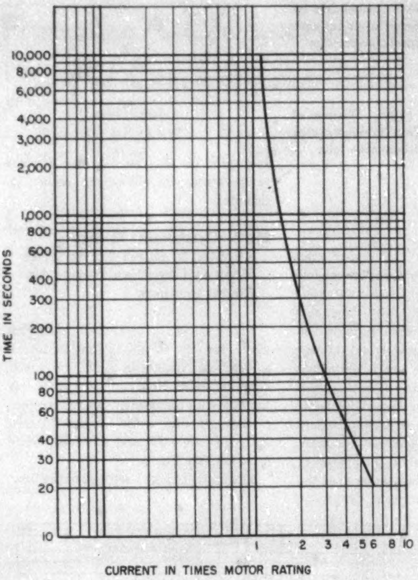


Fig. 4. Time-current (heating) curve for 10- to 500-hp general-purpose induction motors

time-delay and instantaneous trips are adjustable. The time-delay trip is adjustable also. The current pickup settings have ±10% tolerance in accordance with NEMA standards. The operating time band enclosed by the solid lines consists of resettable delay time allowance and the actual tripping time, as indicated on the curves. The resettable delay time line shows the time at which any given magnitude of current must be reduced to 80% or less of the time-delay pickup setting, else the mechanism will continue to move and will eventually trip the breaker. For example, curve *1C*, Fig. 3, shows that a motor drawing six times normal starting current must come up to speed, thereby reducing the current through the circuit breaker trip coil to 80% of its pickup current, in not more than 3.6 seconds, or the circuit breaker will be tripped unnecessarily. This is due partly to mechanical inertia, but mostly to the fact that the air gap between the moving and the stationary armatures has partially closed by that time. This increases the pull exerted by a given amount of current, and it is necessary to reduce the magnitude of the current to compensate for it.

The tripping-time bands (shown in Fig. 3 between the broken lines and the upper solid line of curves *1A*, *1B*, and *1C*) include allowance for variations in performance between successive operations of one particular device as well as variation between devices due to manufacturing variations. A specific breaker will have a relatively narrow tripping-time band, that will lie somewhere within the over-all tripping time tolerance bands shown. If greater accuracy of performance is desired, the particular breaker can be tested to determine its specific characteristics, and the current and time settings adjusted to give tripping characteristics closer to the motor-heating curve.

The time adjustment of the long-time delay mechanisms can be varied between the limits represented by the minimum time setting curves of the *1C* trip and the maximum time setting curves of the *1A* trip; see Fig. 3. When the time setting is changed, the entire operating-time band moves up or down between these limits. In the case of the *1B* and *1A* trip curves, the range through which the lower boundary of the band can be moved is indicated by the short lines intersecting the six-times-normal current ordinates.

Sometimes, circuit-breaker tripping-mechanism characteristics are shown as single lines with a plus-or-minus time tolerance. Such curves are representative of average performance, and the operating-time curve of a specific breaker mechanism might be above or below it.

When reading the curves in Fig. 3, it should be borne in mind that whenever the upper boundary of the tripping-time band is below the motor-heating curve, as shown by the *1C* curve, the mechanism will trip before the time shown on the motor-heating curve. When the motor-heating curve is inside the mechanism's tripping-time band as shown in curve *1B*, the mechanism may or may not operate before the time shown on the motor curve, depending on whether the particular mechanism operates on the lower or upper side of its tripping-time band. When the stalled-rotor-time end of the motor-heating curve is entirely below the tripping mechanism's tripping-time band, as shown by the *1A* curve, the mechanism cannot operate before the permissible stalled-rotor time is exceeded.

All of the curves in Fig. 3 are shown with the current pickup of the circuit breaker set at 125% of motor rating, as specified by The National Electric Code for 115% service-factor motors. However, the circuit-breaker mechanism meets NEMA standard requirements, if it picks up anywhere within its ±10% tolerance band. Therefore a standard tripping mechanism may operate anywhere between the range of 112% and 138% of motor rating.

The *1C* curve, with the minimum-time setting shown could be used for a quick-starting motor; i.e., one that gets up to speed in 3.8 seconds or less, as indicated by the dot at six times normal current on the resettable-delay time line. With the setting shown, the breaker would trip before stalled-rotor time was reached, because the upper side of the tripping-time band is well below the locked-rotor-time end of the motor curve. The latter would not be true, however, if the mech-

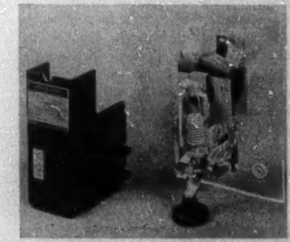


Fig. 5. EC-2 long-time-delay and instantaneous direct-acting trip device for AK circuit breakers

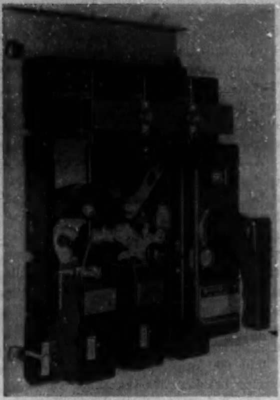


Fig. 6. AK-1-25 circuit breaker with EC-2 tripping device

anism's time setting were changed to the maximum position to allow for a longer starting period. It is apparent that the mechanism overprotects the motor by allowing very short time on moderate overloads. Such a setting might be satisfactory for a motor drive that is unlikely to have overloads, or that has them for very short durations.

The *1B* characteristic curve as drawn shows the mechanism's performance when set to allow 12 seconds at six times normal current, as indicated by the dot on the resettable-delay time line. This combination allows considerably longer time for moderate overloads, but it might not operate before the stalled-rotor time allowance is exceeded, since that end of the motor-heating curve is at the lower side of the tripping-time tolerance band.

The *1A* characteristic curve is shown on its maximum-time setting allowing 40 seconds' motor-acceleration time at six times normal. This curve shows that the permissible heating time shown by the motor-heating curve would be exceeded even in the low-current range if the mechanism should operate on the upper side of its tripping-time zone. The stalled-rotor time would be exceeded, even if the *1A* mechanism were set at its minimum time adjustment, because that end of the motor curve would still be below the tripping-time band.

Short-Circuit Protection Afforded by Series Trips on Large Low-Voltage Air Circuit Breakers

The instantaneous trips on the *EC-2* mechanism are available in ranges from four to nine, or nine to 15 the times breaker rating. The 9-to-15 times rating range is the one to use for motor circuits, because the trip must be set high enough to prevent false tripping during motor starting, or when the motor is pumping short-circuit current into a fault elsewhere on the system. The pumpback current, which is essentially the same as the locked-rotor starting current, usually lasts not more than 1 or 2 cycles, but that is long enough for the mechanism to operate. It is possible for both the starting current and the short-circuit pumpback current to be assymmetrical, i.e., to have a d-c component. For low-voltage motors, (600 volts and under) this d-c component may be sufficient to increase the rms current in the most offset phase to approximately 1.5 times the symmetrical value. The tripping mechanism must be set high enough to allow for this.

Incidentally, it should be noted that the d-c offset multiplier used for calculating the expected current for operation of circuit-breaker tripping mechanisms and relays is 1.5, whereas the multiplier adopted by NEMA for calculating the available current for determining circuit-breaker short-circuit duty is 1.25. The latter gives the average asymmetrical current in the three phases, whereas the tripping mechanism sees the full amount of current in the most offset phase.

If the locked-rotor starting current is six times normal, the asymmetrical rms value in the most offset phase would be $6\times1.5=9$ times motor-rated current. Since the instantaneous trip on the circuit-breaker mechanism has a ±10% pickup tolerance, it would be necessary to set the tripping mechanism at 9.9 times motor-rated current, plus a safety factor to prevent unnecessary tripping during starting, or short-circuit pumpback. The safety factor allows for: 1. possible increase in starting current due to high voltage, and 2. inaccuracies in setting the device. The curves in Fig. 3 show an instantaneous pickup setting of 11.5 times motor rating, which should be high enough to prevent tripping during motor starting, or short-circuit pumpback. If the instantaneous trip is not adjustable, the usual practice is to supply a 15 times rated current trip for breakers up to 800 amperes, and 12 times normal from 800 to 1,600 amperes.

The instantaneous-trip setting should be kept as low as possible, in order to minimize: 1. the damage to the motor in case of an internal short circuit, and 2. the disturbance to the rest of the system by disconnecting a faulted motor quickly. The importance of keeping it low is demonstrated by the long tripping times shown by the curves in Fig. 3 for the relatively high values of short-circuit current just below the values needed for instantaneous tripping.

It would be possible to provide supple-

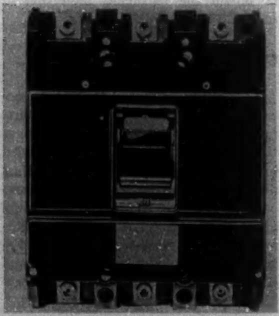


Fig. 7 (left). J-frame, 225-ampere, molded-case circuit breaker

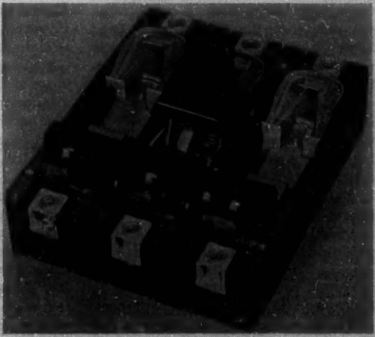


Fig. 8 (right). Interior of J-frame, molded-case circuit breaker

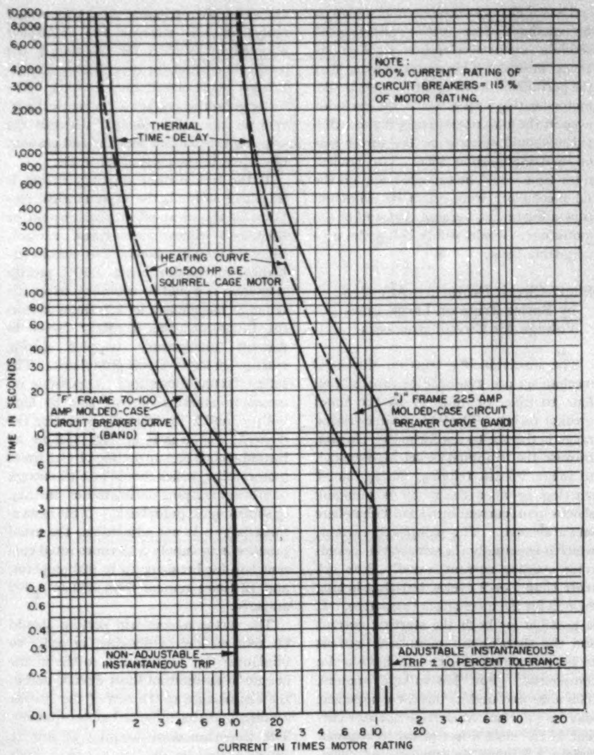


Fig. 9. Time-current curves of induction motors and thermal-magnetic molded-case circuit breakers

mentary induction overcurrent time-delay relays and set them to operate in less than one second on short-circuit currents between locked-rotor starting current and the setting of the instantaneous trip. However, it would not be very much worth while because it speeds up the tripping operation in such a narrow range of current. Incidentally, an induction relay is not affected by the d-c component, which accounts for the fact that it can be set down close to the motor-starting current.

Motor-Overheating Protection Afforded by Series Trips on Molded-Case Circuit Breakers

Molded-case circuit breakers are generally used in combination with conventional motor starters. In such installations, the circuit breakers provide motor-branch circuit protection, and the thermal relay in the starter provides running overload protection for the motor. However, they are sometimes used alone as motor starters.

Figs. 7 and 8 show external and internal views of the General Electric 225-ampere *J*-frame molded-case air circuit breaker. This particular model has a nonadjustable, indirect-heat, bimetallic, thermal time-delay trip, and an adjustable instantaneous trip. Other types and ratings of molded-case circuit breakers have instantaneous magnetic trip only, or a combination of nonadjustable thermal and instantaneous magnetic trip.

Fig. 9 shows the motor-heating curve of Fig. 4 combined with the time-current curves of a 70- to 100-ampere *F*-frame breaker having nonadjustable instantaneous trip and a 225-ampere *J*-frame breaker with adjustable instaneous trip. Both circuit breakers have nonadjustable, indirect-heat, thermal time-delay trips. The circuit-breaker time-current curves are drawn for a 115% service-factor motor with one times breaker rating corresponding to 115% of motor rating.

As shown by the dots on the lower side of their operating-time bands, the time-delay tripping mechanism on the *F*-frame breaker would allow a motor-starting time of 7 seconds with current equal to six times breaker rating, and the *J*-frame tripping mechanism allows 8 seconds motor-acceleration time. The curves should be interpreted in the same manner as the large air circuit-breaker tripping mechanism curves already discussed.

The 600-ampere *L*-frame molded-case circuit breaker has the same curve as in Fig. 9, but is a little slower throughout most of the thermal-trip zone.

The operating-time tolerance bands shown are broad enough to include the manufacturing variations between standard-production, molded-case circuit breakers. The band for an individual circuit breaker will be substantially narrower. Each circuit breaker can be specially calibrated at the factory to determine its own operating-time band, if desired and specified.

Short-Circuit Protection Afforded by Molded-Case Circuit Breakers

The instantaneous tripping mechanism on both the *F*- and the *J*-frame circuit breakers can be set at a maximum of ten times normal; hence it would be satisfactory for use with motors requiring starting current of six times normal or less. In those cases where the starting current is less than six times normal, it would be advisable to set the adjustable instantaneous trip on the *J*-frame units at less than their maximum of ten times normal, to obtain the maximum protection in case of a short circuit.

Conclusions

This study indicates that fairly accurate information as to: 1. motor-starting (locked-rotor) current, 2. starting-time, and 3. stalled-rotor time, is needed in order intelligently to apply and adjust direct-acting series-tripping mechanisms on low-voltage circuit breakers. It would facilitate their co-ordination if the motor-heating curves were plotted on the 4½ ×5 cycle, log-log paper (Keuffel and Esser sheet *336E*) used for plotting time-current curves of the protective devices.

It was found that there are no industry standards regarding permissible motor stalled-rotor time. Such standardization would be very desirable, in order that a given time-current setting of a protective device would afford the same degree of protection for motors of different manufacture, although designed for the same loads.

An examination of the curves in Fig. 3 indicates that the direct-acting series trips on large low-voltage air-circuit breakers, as manufactured, afford a considerable degree of protection to motors that start easily and are not likely to have any severe overloads. The possibility of the circuit breaker's tripping mechanism operating on stalled-rotor condition, when set high enough to prevent tripping during the starting period diminishes until it eventually becomes impossible for it to do so, as curve *1A* shows. In such cases, it would be possible to use supplementary equipment for locked-rotor protection and to let the circuit breaker provide such protection on moderate overloads, as curve *1A* indicates.

The large air circuit-breaker tripping mechanisms will give better performance if tested in the field to determine the width of their operating-time bands and then adjusted to make it come as close as possible to the motor-heating curve.

It is possible to have one of the three tripping mechanisms on the large air circuit breakers sound an alarm on currents close to normal full load instead of tripping the breaker. The use of such a device is debatable on most drives, however, as it is likely to become a nuisance by sounding too often.

Since their time delay is nonadjustable, the use of molded-case air circuit breakers as motor starters is limited to motors that will start within the time indicated by the lower boundary of their time-delay curves for a given value of starting current, both the *F*- and *J*-frame mechanisms would underprotect their motors in the moderate overload range. This should not be troublesome, however, if the circuit-breaker application is limited to installations where the possibility of overloads occurring is small, since the temperatures resulting from the overloads permitted would not damage the insulation unless frequently repeated.

A molded-case circuit breaker with a characteristic similar to the *F*-frame unit would provide adequate protection against stalled rotors. A *J*- or *L*-frame circuit breaker, however, might permit dangerous overheating in case of a stalled motor, with heating characteristic as shown in Fig. 4, if the particular circuit breaker was one that operated on the upper side of its tolerance band. If desired, molded-case circuit breakers can be obtained with special calibration, thereby obtaining a much narrower time-delay operating-time band which would facilitate matching circuit-breaker and motor-heating curves.

In those cases where a circuit breaker's tripping mechanism might be expected to give acceptable protection for ordinary overloads, but not for locked-rotor condition, the absence of locked-rotor protection might be condoned if the motor is in sight of the operator, and the nature of the drive is such that he can visually detect a stalled-rotor condition.

Digital Telemetering and Print Out of Pipe-Line Information

MAX T. NIGH
NONMEMBER AIEE

THE trend toward closed-system pipe lines and remote control has brought many new improvements in the pipe line industry. However, with these innovations have come problems. Remote operation of pipe-line machinery necessitates bringing much information to a controlling point from the "satellite" stations. Transmission and display of this information for best utilization is a subject of interest to all pipe line engineers.

This paper describes a new method of telemetering using supervisory or centralized control equipment. Attention is directed specifically to the how and why of this new method. Digital telemetering offers considerable promise in terms of accuracy, reliability of data transmission, minimum communication-channel requirements, and ease of utilization of data in retransmission or computation.

Development

In the past, the most common method of telemetering pipe-line information has been the time-impulse method. This method involves keying a narrow communication channel for a length of time proportional to the amount of the measured variable. The on and off time periods are complements of each other and establish a definite proportional position of a pointer at the dispatching station. Another system varies the frequency of a signal imposed on the communication channel proportional to the measured variable. These methods, in effect, are analog methods as opposed to digital methods of transmission. To better define the terms analog and digital one may compare a slide rule, which is an analog machine, to a Monroe or Marchant calculator, which, in effect, is a digital or coding machine. The slide rule is approximate, the calculator is exact within the limits of its structure. The centralized control system under discussion is essentially a digital system.

With the analog type of telemetering system, the accuracy of readings obtained is dependent upon the uninterrupted pulsing of the communication channel. Noise and signal interruption of any communication system, be it microwave or wire-line, can contribute to inaccuracies in the measured result at the dispatch point. This may result in "wiping" or oscillation of the pen on an analog telemetering recorder, which, while recognized, gives concern to the dispatcher in observing changes in the measured variables.

This is in contrast to the security against false registry of information, designed into a centralized control system, which transmits the supervisory indications from the satellite station to the controlling or dispatch stations in some form of code steps. In supervisory control systems, the indications can be secured by any of the well-known methods such as use of redundancy, three out of five check, and/or message totalizing. If incorrectly

---

Paper 57-38, recommended by the AIEE Petroleum Industry Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 25, 1956; made available for printing October 27, 1956.

MAX T. NIGH is with the Service Pipe Line Company, Tulsa, Okla.

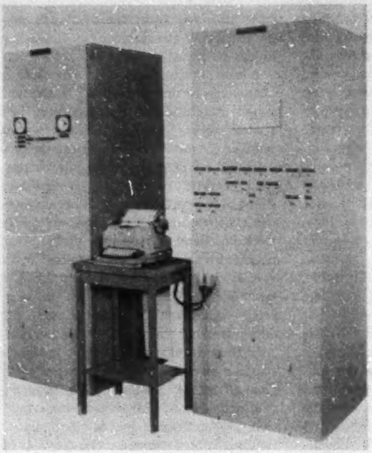


Fig. 1 (left). Office and satellite-station centralized control cabinets

Fig. 3 (right). Telemetering transmitter and encoder

received, the system can be arranged to drop out the data before registry and ask for a repeat until the data is obtained correctly. This security of data against transmission error is very desirable and is necessary in a supervisory control system, as the history of the art has shown. A review of the principles involved in the supervisory system discussed here are given in more detail by Eckhardt.[1]

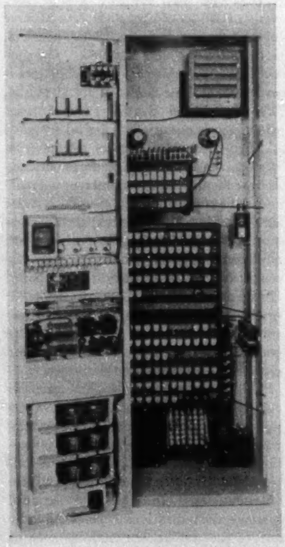


Fig. 2. Rear-view satellite-station centralized control cabinet

The approach to a supervisory or centralized control system to increase its work load for telemetering necessitates consideration of the duty time in pipe-line operations. The scheduling of oil movements in pipe lines is usually made on a weekly or monthly basis and, therefore, most machinery operations are made somewhat rigid for a similar or longer period. This is true whether the machinery is locally or remotely controlled. A corollary of this might be that the machinery has very few breakdowns or losses of power during such periods, if properly engineered and maintained, and therefore few stops and starts are made. It follows then that a standard supervisory control system, in pipe-line service, gets very little use in comparison with its capacity for work, as compared with the work obtained from most machinery in a pipe-line system. With these points as background, it is easy to see why the incentive is created to make better use of the exceptional tool of supervisory control.

The equipment described in the following is designed to reach the objective of using supervisory control equipment for telemetering as well as for remote control.

Design

The digital-telemetering idea was first applied by the Service Pipe Line Company to a single 1,250-horsepower booster station on a 12-inch line connecting two stations with tankage in central Wyoming. The local facilities and control equipment of this electric centrifugal station are considered to be good standard design in the pipe-line industry and are not discussed in this paper. An excellent review of these details may be obtained in the article by Anderson.[2] The communication link is 25 miles of telephone line, with all communication channels for control and supervisory indications above 37-kc frequency.

The centralized control system for the satellite station provided for complete remote operation of the pumping unit. The following indications of machinery status are transmitted back to the controlling station:

1. Pumping unit on.
2. Pumping unit off.
3. Suction valve open.
4. Suction valve closed.
5. Main circuit breaker open.
6. Low suction pressure.
7. Communication-circuit failure.
8. 50-psi (pounds per square inch) low-air-pressure station warning.
9. Pump seal failure or high sump.
10. Current unbalance or undervoltage.
11. High pump case pressure or temperature, or high discharge pressure.
12. Pumping unit fault, e.g., motor overload, motor bearing, motor-winding temperature, motor lube-oil pressure, trip battery, low voltage, pump vibration, 20-psi air-pressure alarm.

Note that certain faults were grouped together to conserve cost in centralized control equipment and still provide all the necessary information for the most efficient type of operation from the controlling station.

Service Pipe Line Company has put a number of stations similar to the satellite station described in service and have used time-impulse telemetering of suction, case and discharge pressure, and flow to the controlling station. With this analog system, six communication channels have been required, one for control codes, one for indication codes, and one each for the four measured variables. In some of these installations is was necessary to use signal channels above 37 kc because of the number of telephone carrier channels occupied by voice communications. This made the requirement for repeaters necessary where any great distance was involved for the transmission. When several stations are involved, or where leased wire facilities are the means of communication, the advantage of reducing the number of communication channels is obvious.

At this satellite station the number of communication channels were reduced from six to two, by telemetering all the information of the measured variables on the same signal channel used for supervisory indications. The programmer at the satellite station transmits each measured variable sequentially. At the controlling station all these variables are stored and then printed out in sequence on the typewriter, with the time at which the telemetered data was requested.

Although only four measured variables are transmitted from the satellite station, a much greater number of measurements could be sent on the same channel with a small increase in the size of the programmer.

Fig. 4. Typical print-out record of relay-station starting

Deviation settings: flow, 80 b/h, pressure, 50 psi
Rate of change: flow, 70 b/h per minute with 40 b/h threshold, pressure, 25 psi per minute with 20 psi threshold

```
20:55    S 610    C 610    D 610    F 1940
21:10    S 610    C 610    D 610    F 1940
21:25    S 610    C 610    D 610    F 1940
21:40    S 610    C 610    D 610    F 1940
21:55    S 610    C 610    D 610    F 1950
22:06    S 570    C 940    D 620    F 1970
22:07    S 450    C 720    D 680    F 2180
22:08    S 310    C 670    D 670    F 2300
22:09    S 250    C 650    D 650    F 2210
22:10    S 200    C 640    D 640    F 2140
22:13    S 150    C 630    D 620    F 2050
22:14    S 150    C 650    D 620    F 1990
22:15    S 160    C 650    D 610    F 1980
22:16    S 160    C 650    D 620    F 1980
22:31    S 160    C 660    D 620    F 1970
22:47    S 160    C 660    D 620    F 1970
23:02    S 160    C 660    D 620    F 1960
23:16    S 160    C 650    D 620    F 1960
23:31    S 160    C 650    D 620    F 1960
23:47    S 160    C 650    D 620    F 1960
```

Equipment

Fig. 1 shows the control cabinet located at the satellite station, the typewriter, and the control cabinet located at the controlling station. Fig. 2 shows the rear of the control cabinet located at the satellite station. On the swinging panel are shown the programming-timer and channel-deriving units. Inside the cabinet from top to bottom, are the heater, telemeter transmitters, coding unit, storage and programming relays. In the bottom box, the deviation and rate of change relays for discharge pressure and flow are shown.

The telemeter transmitters each consist of a pressure-measuring system and an encoder. The pressure-measuring system consists of a null-balance servo-operated indicator with a differential transformer in the transmitter being positioned by the bourdon tube or bellows. The servomotor is driven by the amplified error signal between the differential transformer in the transmitter and the differential transformer in the receiver. The encoder is a device that converts a shaft position into a digital output, in our case the indicator pointer shaft. This conversion is accomplished by 12 brushes that ride on a code plate with alternate silvered and insulated spaces on its surface, establishing a binary-code combination. This characterizes the measured variable in digital form. At the satellite station all measured variables are made up of three significant figures, and thus 12 output circuits are required to supply the binary bits to signify numerical values from 0 to 999. Four of the bits represent the thousands values, four bits represent the hundreds values, and four bits represent the tens values. Pressure and flow measurements could not be read closer than to the nearest 10 pounds and therefore closer values are not transmitted. This is a very important design point. The greatest economy of transmission is obtained when no value is transmitted at a greater resolution than is possible to read at the satellite station. Measurements at the satellite station are made as accurately as possible. The accuracy obtainable with the telemetering-measuring system is guarded by the code-transmission system, and thus the registry of error is eliminated in the telemetering system. Fig. 3 shows the telemeter transmitter consisting of (from left to right) the encoder, servomotor, and commutator (discussed later), and the measuring-system receiver or indicator. The differential-transformer transmitter is not shown, nor is the power amplifier for the servomotor. The servomotor shown in Fig. 3 receives its power from a magnetic amplifier measuring the error signal between the transmitter differential transformer and the receiver differential transformer. Industry in general is showing considerable interest in instrumentation of this type. It offers features which conventional air or electronic instrumentation do not possess, i.e., null-balance design with considerable power output to operate a number of devices, and a circuit involving no vacuum tubes. A code plate is the most convenient method of obtaining output from a shaft position, and an instrument with power output is required to position the encoder.

Transmission of these digital data is accomplished in the same manner as all supervisory indications. The measured variable through the encoder is stored in 12 relay positions in the supervisory cabinet at the satellite station. These relay positions are transmitted as long and short pulses to characterize their position and at the office or controlling station the long and short pulses determine the position that certain relays, designated as storage, are to assume. When the code is received into the controlling station it is stored in weighted-decimal binary form. This means that the four bits which represent each digit are assigned weighted values such as 1, 2, 4, and 8. Various combinations of these four digits can determine any number between 1 and 10. A programmer now sweeps past these data in storage and thus determines which keys to operate on the typewriter. Fig. 4 shows a typical print-out record for the starting of a pipe-line booster station.

Four methods determine the frequency of sending this information back to the controlling station:

1. An adjustable timer (0 to 60 minutes) in the satellite-station cabinet.
2. The dispatcher at the controlling station asking for information by pushing a button.
3. A counting circuit measuring the deviation from the previous reading; this is adjustable from 40 to 100 psi.
4. A rate-of-change device which not only sends in the telemetered data but asks the typewriter to print the data in red as long as the upset continues and then in black for every minute for 5 minutes after the upset ceases.

The time is also determined digitally at the controlling station and is printed on the typewriter at the start of each line of data. The time consumed between requesting data and obtaining a full line of information is approximately 20 to 30 seconds. A typical example of this type digital display is shown in Fig. 4.

Special Features

As mentioned in the foregoing, the controlling station can request information or a timer can automatically send in information. Another device at the satellite station can initiate the transmission of operating information. The "silent dispatcher" is applied to the two measured variables, discharge pressure and flow rate. This monitoring equipment is of the computer type, and consists of relays and the commutator with brushes. The transducer for it is the incremental coder (commutator) shown in Fig. 3. Each bar represents 10-b/h (barrels per hour) or 10-psi change in flow or discharge pressure. The computer receives its intelligence from four brushes (two are common) riding on this commutator. Relays count each 2.5-unit change from the previous reading used as a reference point, and, with deviations either up-scale or down-scale, this device adds the units of deviation into the counter. The counter adds in units of 10. As the measured variable reverses itself and starts to return toward the reference point, this device subtracts the units of deviation from the counter. Thus, the ordinary small fluctuating conditions which are quite natural in pipe-line operations never "rack up" sufficient count to initiate a print-out request. However, when the deviation continues in the same direction and adds up to the given preset amount (adjustable from 40 to 100 psi or b/h) in 10-unit increments, this device recognizes the excessive deviation and sends a complete set of telemetered data to the controlling station. Having reported a set of data, a new reference point is established, from which new deviations are monitored.

Another part of this computer is a rate-of-change unit consisting of a timing device to sense whether these deviations occur at a slow or fast rate. In batching operations, large amounts of deviation may be expected but the rate will be very slow. However, rapid deviation may result from an upset in the system, e.g., a line break or a gate change at the receiving point, and should be reported immediately. The alarm signal is adjustable from 1-b/h or 1-psi-per-minute to 30-b/h or 30-psi-per-minute rate of change. A threshold deviation of 20 b/h or 20 psi must be obtained before the rate of change will cause an alarm. This device initiates the transmission of a complete set of data to the controlling station. The typewriter prints the data in red and an alarm sounds to direct the operator's attention to the upset. In Fig. 4, the lines printed in red are indicated by larger type. Additional printing of data in black is sent in every minute for 5 minutes after the upset ceases to establish the new operating condition.

A significant point about the silent-dispatcher concept is that a "key" measured variable is being monitored continuously. The dispatcher at the controlling station can look at the last line of data printed and know that the operating conditions at the satellite station have not changed appreciably since the time it was printed, providing the communication circuit is intact. Monitoring of the communication channel is usually a standard feature of supervisory control systems. The need for continuous readings or constant observation on the part of the dispatcher are eliminated.

A deviation-type device is necessary in pipe-line operations where batching of crudes of different viscosity and variations in flow rates are constantly changing the operating pressures. A pressure switch set to operate close to the normal discharge pressure could approximate the action of the deviation device if, and only if, conditions never changed. Most pipe-line operations are certainly not this constant.

Future System Application

The techniques discussed in the foregoing represent a solution to many of the monitoring problems attendant with single or multiple-pipe-line operation, involving a number of satellite stations where batching of crudes is practiced, and where varying rates of flow are encountered. The principles discussed in this paper are being applied to other multiple-station remote operation on Service Pipe Line Company's system. When the number of variables to be used in pipe-line monitoring reach 12 to 60, the transmission of this much data and the method of displaying the information for dispatching use become problems. The digital-type display shown in Fig. 4 is most easily assimilated by the dispatcher and certainly makes his task much easier.

On systems which are being installed on our pipe-line system involving multiple stations, a teletype-code output is also provided at the controlling station to permit retransmission of the data to a central point and can be used in programming into a computer when desired.

References

1. TIME CODE CONTROL AND DATA HANDLING SYSTEMS, P. K. Eckhardt. *Petroleum Engineer*, Dallas, Tex., Aug. 1956, pp. D-29, D-32.

2. MIDGET BOOSTER STATIONS DRAWING BIG JOB, L. E. Anderson. *Pipe Line Industry*, Houston, Tex., Aug. 1956, vol. 5, no. 2, pp. 46–49.

◆

Discussion

W. A. Derr (Westinghouse Electric Corporation, East Pittsburgh, Pa.): Mr. Nigh's paper is of great interest since local and remote automatic logging of the magnitudes of important quantities is being considered for stations in many industries.

The part of the installation made by the Service Pipe Line Company, which concerns the magnitude of variable quantities, may be considered as three separate and basic systems, as follows:

1. Telemetering.
2. Automatic data logging.
3. Continuous monitoring of high and low conditions of the quantities logged.

The telemetering is only required because the application described concerns a remote station. For local-station applications, automatic data logging and continuous monitoring can be provided without the necessity of telemetering.

Selective telemetering of quantities in a remotely controlled station dates back to the early days of supervisory control in the 1920's. Until the rather recent advent of printed records of telemetered quantities, such selective telemetering was usually done by analog methods. The channel requirements for such selective analog telemetering are the minimum possible, since such telemetering can be accomplished over the supervisory control channel. Unless a printed digital record is to be provided, such as described in Mr. Nigh's paper, selective telemetering on an analog basis will usually be preferred over digital telemetering because of its simplicity and considerably lower cost. For selective telemetering, inaccuracies caused by channel conditions will not occur for analog-type telemetering any more than for digital-type telemetering.

When the telemetered quantities are to be printed in digital form, such as in the system described, the telemetering can be done on an analog or a digital basis. For such a logging system, the information concerning the magnitude of each telemetered quantity must be fed to the printing equipment in digital form. In the system described by Mr. Nigh, the conversion to digital form was made at the remotely controlled station, with series of pulses being transmitted to the controlling station to designate the magnitude of each digit. It is possible to telemeter in analog form and to convert to digital form at the location of the printer. The detailed method employed must be governed by economic considerations and desired operating procedures and accuracies for a particular installation.

Mr. Nigh's paper discusses the continuous monitoring of high and low conditions of the telemetered quantities. While continuous monitoring for record purposes will often be desirable or even necessary, there will be installations in some industries where such monitoring for record purposes cannot be economically justified. On an installation such as described by Mr. Nigh, a considerable reduction in equipment would result if the continuous monitoring feature were eliminated. A common servo and encoder could then operate in conjunction with an unlimited number of transducers. The programmer would then connect the servo to each transducer in sequence. This would increase the readout time but there will undoubtedly be many applications where this will be permissible.

Selection of Buffer Reactors and Synchronous Condensers on Power Systems Supplying Arc-Furnace Loads

C. CONCORDIA
FELLOW AIEE

L. G. LEVOY
FELLOW AIEE

C. H. THOMAS
MEMBER AIEE

Synopsis: The studies and analyses presented in this paper provide a rational method for the selection of the optimum buffer reactance and synchronous condenser size to limit flicker-voltage disturbances caused by arc-furnace loads on power systems in those cases where such corrective measures are required. The influence of the system parameters, including buffer reactance, condenser size, system reactance, and amount of additional steady loads such as furnaces in the refining stage of the melt, are considered simultaneously with shock loading produced by a furnace operating in the melt-down stage. In addition, the magnitude and frequency of the shock load are varied over a considerable range, since there is no generally accepted typical shock.

MOST arc furnaces are served directly from power systems of adequate size and stiffness to limit flicker voltages to acceptable values. Sometimes the furnace load can be isolated from other loads by serving it directly from a point in the system which has adequate stiffness. However, occasionally it is necessary to provide some means of alleviating the flicker voltage caused by such furnaces, and the synchronous condenser with buffer reactor provides one means of accomplishing this.

The increasing use of electric arc furnaces in the steel industry, the increasing size of such furnaces, and the requirements by utility companies to limit voltage disturbances on their systems in the interests of improved quality of service to customers, make it desirable to establish criteria for the selection of the most economical and satisfactory combinations of synchronous condenser and buffer reactor for any case where such correction is needed.

The arc-furnace load is characterized by rapidly fluctuating kilowatts and kilovars during the early phases of the melt down period. The rapid kilovar changes are of particular concern because they tend to cause voltage changes in the supply and may create a flicker voltage problem if the supply circuits are not stiff enough.[1,2,3]

The use of a synchronous condenser and buffer reactor partially to shield the utility system from the rapidly changing kilovar load is illustrated in Fig. 1. This combination is very effective if properly sized and proportioned.

Since the buffer reactance is considerably less expensive than the synchronous condenser, it is desirable to use as large a buffer reactance and as small a synchronous condenser as will meet the requirements.

In any given installation, as the buffer reactance is increased, it is shown in this paper that the voltage disturbance on the critical bus for a given shock load first decreases, reaches a minimum, and then increases. This is because the power shock that accompanies the reactive shock produces a transient swing of the condenser angle. When the buffer reactance is so large that the synchronous condenser swings through a large angle with respect to the system voltage, eventually a point is reached where the swinging of the condenser aggravates the disturbance on the system, and a further increase in the buffer reactance results in poorer flicker performance on the critical bus. One of the purposes of this paper is to explore this limitation in buffer reactance.

One previous study[4] had indicated that the buffer-plus-system reactance should be limited to about 35% on the condenser base for stability. Another[5] has concluded that for stable operation of one arc furnace the buffer reactance should not exceed the arc-furnace reactance. The present paper extends this last study by taking into account the effects of having more than one furnace and of more than one shock load.

The analysis presented in this paper shows that stability in the usual sense is not the first limitation reached as the buffer reactance is increased. Instead,

---

Paper 57-56, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 22, 1956; made available for printing November 29, 1956.

C. CONCORDIA and L. G. LEVOY are with the General Electric Company, Schenectady, N. Y., and C. H. THOMAS is with the University of Kansas, Lawrence, Kans.

The authors wish to express their appreciation to the Northeastern Steel Corporation and the United Illuminating Company for providing the opportunity to make the tests shown in Fig. 4.

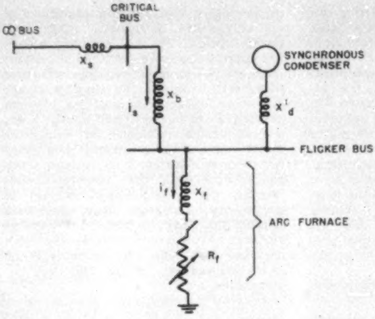


Fig. 1. Simplified one-line diagram of arc-furnace power supply showing relation of synchronous condenser, buffer reactor, and arc furnaces

the swinging of the condenser with respect to the system will eventually reach a magnitude where the voltage flicker on the critical bus becomes more severe under the shock loading of the furnace than would be the case with a smaller reactor. The analysis also indicates that it is practically impossible to force the condenser out of step by either the application or removal of a single shock load of the maximum possible magnitude in a typical arc furnace.

Stability Considerations

There are two types of disturbances that can be considered under the general heading of stability for the system shown in Fig. 1.

1. A fault in the supply system depressing the supply voltage for an interval of time until the fault is cleared by appropriate relaying and circuit breakers thus restoring normal voltage. Such a disturbance, if sufficiently prolonged and severe, can cause the synchronous condenser to fall out of step with the system. An examination of the system parameters and establishment of the limiting conditions for satisfactory performance under such fault conditions and under normal power transfer conditions falls in the area normally encompassed by a stability study, and the term "instability" is usually used to describe the situation where, on restoration of voltage and normal circuit conditions, the machines have pulled too far apart during the disturbance to permit them to pull back successfully into synchronism.
2. When considering flicker voltage on the critical bus the studies reported in this paper indicate that as the buffer reactance is increased for a given pattern of shock loading, eventually a point will be reached where the disturbance of the critical bus voltage will increase. However, the condenser will not fall out of step at this point where the disturbance increases and consequently the system is not unstable. For this reason the term "optimum" is used for describing the buffer reactance for minimum voltage disturbance on the critical bus under a given pattern of shock loading condition. The point of instability is reached eventually when repeated shock loads of sufficient magnitude and proper frequency are applied to cause the condenser to fall out of step with the system. However, before this point is reached the voltage disturbance on the critical bus becomes excessive, and consequently an out-of-step, or instability, criterion is of less value than is the optimum value in the design of such systems.

Voltage Dip Neglecting Condenser Swings

A one-line diagram of the system is given in Fig. 1. The synchronous condenser is connected to the flicker bus, and the flicker bus is connected through the buffer reactance to the critical bus.

It is of interest first to calculate the voltage dip ΔV_{co} produced at the critical bus without any corrective means. (In this discussion of voltage dip produced by short circuit there is no intention of superseding the common practice of calculating a typical voltage dip on the basis of a typical kilovar swing appropriate to the furnace transformer under consideration. Such typical kilovar swings are found, for example, in references 1 and 3. The shocks considered in this paper are in general much more severe, but are also much less frequent.) In going suddenly from open-circuit to 3-phase short circuit of the furnace electrodes the voltage dip is, neglecting the residual short-circuit resistance:

$$\Delta V_{co}=\frac{X_s}{X_f+X_s}$$

or

$$\frac{X_f+X_s}{X_s}\Delta V_{co}=1.0 \tag{1}$$

Equation 1 suggests that voltage dips ΔV_c obtained under various conditions employing corrective equipment may be expressed in per unit of the voltage dip for a 3-phase short circuit without correction by means of the quantity:

$$\frac{X_f+X_s}{X_s}\Delta V_c$$

With a buffer reactor X_b and a synchronous condenser having transient reactance X_d', as in Fig. 1, the voltage dip produced at the critical bus in going from open circuit to short circuit of the furnace electrodes is given by

$$\frac{X_f+X_s}{X_s}\Delta V_{co}=\frac{\dfrac{X_f+X_s}{X_f}}{1+\dfrac{X_b+X_s}{X_f}\left(1+\dfrac{X_f}{X_d'}\right)} \tag{2}$$

The right-hand side of equation 2 is the ratio of the voltage dip with correction to the voltage dip without correction, for the 3-phase short-circuit case, and is thus a quick though approximate measure of the improvement in voltage flicker realized in a given installation by the corrective means.

In equation 2 the significant parameters are the ratio of buffer-plus-system reactance to furnace reactance $(X_b+X_s)/X_f$ and the ratio of furnace reactance to condenser transient reactance X_f/X_d'. The numerical results given in this paper are therefore expressed in terms of these ratios.

The presence of refining loads tends to reduce the voltage flicker slightly if no appreciable condenser swinging occurs. It will be seen, however, that for the type of load shock considered in this paper refining loads augment any condenser swinging and thus may result in increased voltage flicker for large disturbances.

The value of X_d' that has been used in these calculations is the average of the rated-voltage and the rated-current values. It is evident that the voltage dip at the critical bus can be minimized by

1. Reducing X_d' (increasing the condenser size and cost).
2. Reducing X_s (stiffening the power system).
3. Increasing X_f (reducing the furnace rating and output).
4. Increasing X_b.

This static-circuit analysis indicates that increasing the buffer reactance with a given size furnace and condenser tends to reduce the shock kilovar loading on the system and to make the synchronous condenser take a greater share of the kilovar swing. Since increasing the buffer reactance is much less expensive than reducing the synchronous condenser reactance there is an urge to increase the buffer reactance as much as possible, thereby permitting the use of the smallest

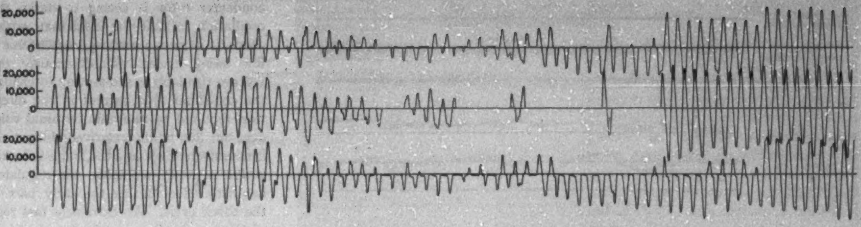


Fig. 2. Three-phase 7,500-kva arc-furnace transformer secondary currents during early part of melt-down on 250-volt tap

synchronous condenser which will give the desired flicker levels at the critical bus and still be adequate to supply the required steady-state kilovars. The study reported in this paper was undertaken to establish rational values for the buffer reactance size for a range of cases.

Synchronous Condenser Rating and Characteristics

The average voltage of the furnace, or flicker, bus must be maintained at rated value to insure rated production from the arc furnaces. To maintain rated average voltage on this bus requires that the condenser kilovar rating be at least equal to the sum of the kilovar demand of all arc furnaces operating from the flicker bus, plus half of the kilovars in the system-plus-buffer reactance. In practice the synchronous condenser rating should be 10% to 20% above the required output because of the roughness of the load. Also, it is frequently necessary to make the condenser rating large enough to accommodate the possible addition of future furnaces. Usually the furnaces are staggered in operation so that the melt-down of one furnace occurs when the other furnaces on the same bus are in the refining phase. However, in some cases melt-down periods can overlap or occur simultaneously. If this is a requirement the condenser must be selected to have adequate kilovar rating to meet the most severe condition of loading which it will encounter.

There may be in addition some intermittent load unbalance and intermittent single-phase operation in an arc furnace, especially during the early portion of the melt-down period. This also contributes to synchronous-condenser heating. However, sustained single-phase operation such as might occur in case of a broken electrode should not be permitted, as such operation may cause serious overheating of the condenser. Normally relays are provided to protect the condenser against sustained operation under such conditions.

Arc-Furnace Load Characteristics

Some of the important features of arc-furnace loads are

1. The maximum kilovar load occurs when the electrodes are short-circuited. This load can be determined from the reactance of the furnace, transformer, and secondary leads and from the operating voltage tap.
2. The maximum kilowatt load for a given voltage tap occurs when the load-circuit resistance and reactance are equal.
3. During the melt-down phase the furnace may be operated at 1.2 times its name-plate rating with a power factor normally in the range of 75 to 85%. During the refining phase furnaces are usually operated at from one fourth to one third of rated power and are relatively steady loads whose power factor is normally in the range of 85 to 90%.

Fig. 2 is a reproduction of a portion of an oscillogram of the electrode currents of a 7,500-kva furnace near the beginning of the melt and shows the erratic nature of the load. Fig. 3 represents a typical oscillogram taken during the refining interval on the furnace of Fig. 2.

A small section of an oscillographic record of another arc furnace[8] (rated 12,500 kva), having a 20,000-kva synchronous condenser and a 49% (on 20,000-kva base) buffer reactor, operating during melt-down is shown in Fig. 4. This section begins 17.6 seconds after the start of the melt-down and shows the effects of disturbances occurring within the next 2 seconds.

The calibrations are given in terms of d-c deflections. Power out of the condenser is shown as a positive deflection. The average power of the condenser is a small negative value representing condenser losses. The single-phase wattmetric elements showing utility power and furnace power give both the kilowatts and kilovars. Note that the utility power factor is close to unity whereas the furnace power factor is considerably lower, especially during disturbances.

The voltage disturbances shown on the flicker bus are quite severe, but the voltage disturbances at the critical bus (not shown) were pratically imperceptible.

From a study of these and other typical performance records it appeared desirable to explore the effect of repeated shock cycles occurring at or near the natural frequency of the condenser swinging on the system, and also to establish the effect of the magnitude of the shock cycle involved.

In this paper a maximum shock cycle is defined as follows:

> When operating at the maximum possible kilowatt load during melt-down, the three furnace electrodes are short-circuited and

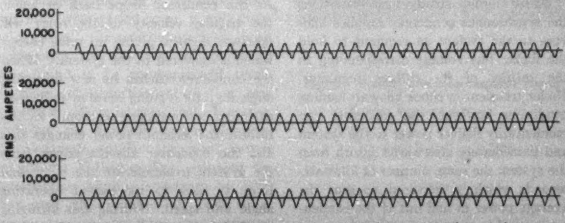


Fig. 3. Three-phase 7,500-kva arc-furnace transformer secondary currents during refining interval on 92-volt tap

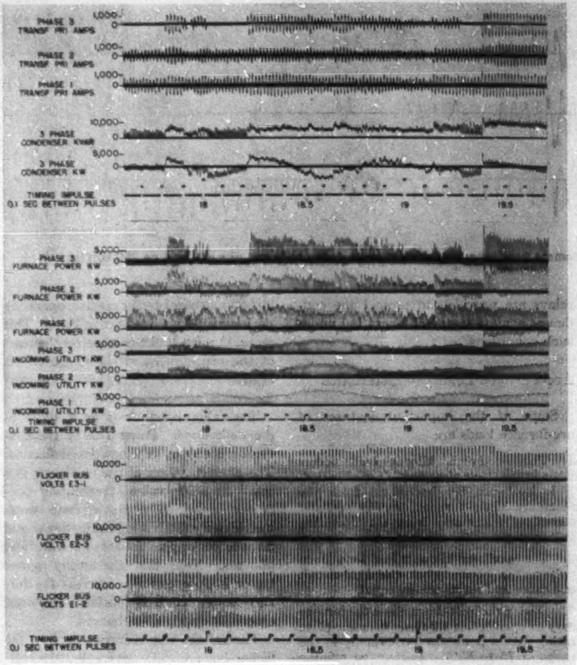


Fig. 4. Performance test on arc furnace, synchronous condenser, and buffer reactor

the load resistance R_f is assumed to go to zero. (In practice, R_f will not go to zero due to losses in the transformer, leads, electrodes, and melt, but this is a conservative assumption since it yields maximum power change.) After the synchronous condenser has reached its maximum excursion, R_f is restored to its maximum load value, thus resulting in maximum overtravel of the condenser.

Analysis

Under normal steady load conditions the synchronous condenser supplies kilovars to the system as required to hold the flicker bus voltage corresponding to the setting of its voltage regulator. Under transient or shock kilowatt loading conditions however, the condenser can momentarily deliver power to the system and immediately afterwards absorb from the system the same number of kilowatt-seconds which it delivered, so that the average power in and out of the condenser is equal to zero (neglecting losses).

To illustrate the action of the condenser in the system shown in Fig. 1, consider what happens if both the kilowatt and kilovar loads of the furnace suddenly increase. The terminal voltage of the condenser (flicker bus voltage) is suddenly reduced and retarded in phase angle as determined by the new loading condition, requiring a change in the operating angle of the synchronous condenser, and an increase in its field strength as signalled by the voltage regulator. As the condenser drops back in angle, the angular velocity of the rotor will decrease slightly, giving up some energy which is delivered to the system. When the condenser reaches its new operating angle its rotor is going too slow so it overshoots its new operating position. At this instant the terminal power changes sign and the condenser absorbs power from the system to accelerate the rotor and bring it back to its correct operating angle and speed. During this swinging of the condenser with respect to the system, the amortisseur winding on the condenser rotor is acting to damp the oscillation so that the condenser settles down at its new operating angle within a few swing cycles. Simultaneously the voltage regulator is attempting to change the condenser field current in the direction required to maintain terminal voltage, but due to the relatively long time constants involved and the limited forcing action normally available, the regulator has little effect during the early part of the shock cycle. An extremely fast regulator may tend to overshoot and add to the disturbance on the critical bus. For this reason practical regulators in use today for such service need not be set for extremely fast action, and normal regulator speeds are satisfactory. In going from open circuit to short circuit in the furnace there is very little change in power, so the condenser does not change its operating angle appreciably. Consequently the static circuit theory (equation 2) yields approximately correct answers for this case.

Examination of actual arc-furnace oscillograms taken at several stages during the melt-down period has indicated that the maximum possible 3-phase power swings were very rare. However, smaller power swings caused by disturbances involving either one or two electrodes were relatively frequent during the melt-down period. These smaller power swings were random in nature and occasionally one would occur at an unfavorable time with respect to the swinging of the synchronous condenser, resulting in a larger excursion of the condenser than is normally associated with a power shock of the magnitude involved.

To permit the rapid exploration of a number of cases and to cover a wide range of system parameters the system shown in Fig. 5 was set up on an analog computer (electronic differential analyzer[7]) using the machine constants and equations given in Appendix I. With this circuit it was possible to study the effects of additional loads and of voltage regulator action.

Principal Cases Studied

The effect on system stability and on magnitude of voltage flicker was determined for various conditions of cycling of the melting arc-furnace load. The principal parameters varied were (1) ratio of system-plus-buffer reactance to furnace reactance, (2) size of the synchronous condenser relative to the melting furnace transformer, (3) per cent refining-furnace load, and (4) magnitude of the load shock.

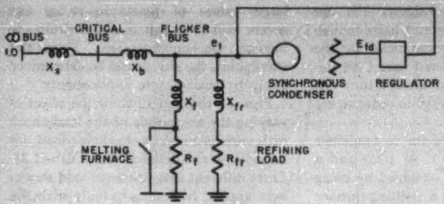


Fig. 5. System studied on analog computer

Data were obtained for some of the foregoing cases at load-cycling frequencies in the range of 0 to 1.8 times the system natural frequency. In particular, most of the results presented here were obtained during the cycling of the melting furnace load at the most unfavorable frequency.

The results were obtained from analog computer response curves as illustrated in Figs. 6(A) and 6(B). The four response curves are: exciter voltage E_f, flicker bus voltage e_t, electrical angle δ of the synchronous condenser with respect to the infinite bus, and the component of voltage drop ($V_c = 1.0 - e_q \cos \delta - e_d \sin \delta$) across the buffer-plus-system reactance in phase with the infinite bus voltage. Since the system reactance X_s is much smaller than the buffer rectance, so that the angle between the infinite-bus voltage and the critical-bus voltage is always very small, the change in V_c is proportional to the voltage flicker at the critical bus. These relations are shown on the vector diagram of Fig. 7.

In the analog computer setup used, alternating voltages were represented by their envelopes. This representation resulted in the d-c offset appearing as a 60-cycle oscillation on the computer records when R_f or portions of R_f were short-circuited. This oscillation is not included in the measurements since it does not affect the voltage flicker.

The method of measuring voltage flicker from the response curves is illustrated in Fig. 8, which shows the critical-bus voltage (proportional to V_c) and the condenser angle following the application of one shock cycle. When the short circuit is applied the voltage dips as shown between the arrows at *a* and the condenser swings through zero angle. This condenser swing produces some very small voltage change as indicated. When the load is restored at *b* the voltage rises again and the condenser starts to swing back through its former operating angle. The voltage rises very slightly as the condenser starts to swing back through its former operating angle. The voltage rises very slightly as the condenser passes zero angle and then dips as shown at *c* or *c'*. As the buffer reactance, or the ratio $(X_b + X_s)/X_f$, is increased the change of voltage produced by this condenser return swing will change from less (*c'*) to more (*c*) than the initial change at *a*. The change at *b* is usually about the same as at *a*. Results are plotted as the largest voltage flicker, whether at *a*, *b*, or *c* (or *c'*), produced by the complete shock cycle. The optimum buffer reactor for the applied shock, to give minimum voltage flicker, occurs when the two flickers (at *a* and *c*) are equal. Note that voltage flicker as defined in this paper is not the same as voltage deviation from normal.

Other Conditions Considered

Certain other secondary effects were considered in the study. For example, a simplified prototype voltage regulator was employed which would be representative of a typical installation. The regulator used was sufficiently slow acting so as not to respond appreciably to condenser swinging and had only a slight adverse effect on system stability under conditions of load cycling. It required approximately 70% as many load shock cycles to force the condenser out of step with the regulator as it did when field flux linkages were maintained constant and the terminal voltage was allowed to drop.

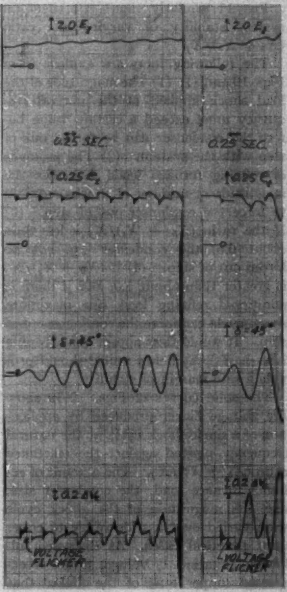
In addition the size of the amortisseur winding was varied over a relatively wide range. Considering one load shock cycle only, the effect on voltage flicker was found to be comparatively small. Increase of amortisseur winding size (decrease of resistance) did, however, improve the over-all system stability. For example, under the most severe conditions of load-cycling, reducing the amortisseur winding resistance by 50% increased the number of load shock cycles required to force the condenser out-of-step by about 50%.

Results

The results of this study are presented in Figs. 9 through 20.

Fig. 9 is a plot of voltage flicker as a function of the frequency of application of the shock for a particular set of system parameters. For this case in the region of the natural frequency of the condenser on the system there is a significant bandwidth where the condenser can be forced out of step by repeated maximum shock cycles. However, even in this region a single maximum shock cycle gives only moderate flicker voltages.

Fig. 10 illustrates the effect on system stability (expressed as the number of shock cycles required to cause the condenser to fall out of step) of both the



(A) (B)

Fig. 6. Analog computer response curves

A—For the case where

$$\frac{X_b+X_s}{X_f}=0.5 \text{ and } \frac{X_f}{X_d'}=1.35$$

B—For the case where

$$\frac{X_b+X_s}{X_f}=1.1 \text{ and } \frac{X_f}{X_d'}=1.35$$

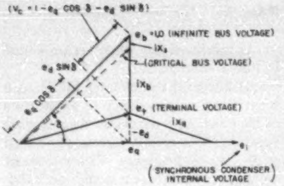

(V_c = 1 − e_q COS δ − e_d SIN δ)
e_b = 1.0 (INFINITE BUS VOLTAGE)
iX_s
(CRITICAL BUS VOLTAGE)
e_d SIN δ
e_q COS δ
iX_b
e_t (TERMINAL VOLTAGE)
iX_d
−δ
−e_d
e_q
e_i
(SYNCHRONOUS CONDENSER INTERNAL VOLTAGE)

Fig. 7. Vector diagram (transient condition, with condenser absorbing power)

magnitude of the load shocks and the amount of refining furnace load. All shock cycles were applied at the condenser natural frequency and at the phase angle to give maximum disturbance. More than one or two such shock cycles would probably never occur in a practical case.

Fig. 11 similarly shows the effect on system stability of varying the ratio $(X_b + X_s)/X_f$.

The following facts are evident from Figs. 10 and 11: (1) The magnitude of the load shock applied at the natural frequency must exceed a certain value before the condenser can be forced out of step with the system. (2) The presence of refining furnace loads decreases the magnitude or number of shocks required to force the condenser out of step. (3) If the ratio $(X_b + X_s)/X_f$ is less than about 0.5, the condenser can not be forced out of step. (4) If $(X_b + X_s)/X_f$ is greater than about 1.0, and if there is additional refining load, one maximum shock cycle can force the condenser out of step. It would thus appear that keeping $(X_b + X_s)/X_f$ below unity[8] should provide a safe margin of stability.

The solid curves of Figs. 12–19 show the voltage flicker produced by application of a single shock cycle at the natural frequency plotted against the reactance ratio $(X_b + X_s)/X_f$. The amount of refining furnace load, the condenser size, and the magnitude of the shock cycle were varied over a wide range to cover most practical cases. As discussed following equation 1, voltage flicker is expressed in per unit of the flicker that would be produced by the same furnace going from open to short circuit without any corrective means and with no additional refining load. The dashed curves shown in the same figures show the voltage flickers calculated neglecting the condenser swing (infinite inertia) and for the case of open to short circuit (where there is no condenser swing). Because of the per-unit voltage used, these open-to short-circuit cases give directly the improvements brought about by the condenser and buffer reactor. For the load shock cases one would have to calculate the per-unit voltage flicker for the corresponding uncorrected case (this uncorrected flicker would always turn out to be smaller than unity) in order to see the improvement.

Figs. 12 and 13 are for a condenser size defined by (X_f/X_d' = 1.35) and a maximum shock cycle obtained by using for each case studied a melting-furnace load resistance given by the relation:

$$R_f = X_f + \frac{(X_b + X_s)X_d'}{X_b + X_s + X_d'} \qquad (3)$$

These figures show both the flicker resulting from the initial short circuit and the flicker resulting from the reapplication of load and subsequent return swing of the condenser, and illustrate that for small values of reactance ratio the first flicker is the larger while for large values of reactance ratio the second flicker is the larger.

In Fig. 12 one might consider the optimum buffer reactance ratio to be either at 0.61 (if only a maximum shock cycle is considered) or at 0.73 (if the open-circuit to short-circuit case is given equal weight). Fig. 13 shows that additional refining furnace load reduces these optimum values. Note that X_f is the reactance of only the melting furnace.

Figs. 14 and 15 show the effect of condenser size. Here the shock load cycle starts from a melting furnace power factor of 0.85. Although this is not quite a maximum shock, it still corresponds to operation of the furnace at about 20% over rating. The condenser size is defined by the ratio X_f/X_d', and the "condenser ratings" are indicated on the figures only to convey a rough idea of the physical size of the condenser relative to one arc furnace. It is seen that at large values of reactance ratio, and severe load shocks, it is not desirable to increase the condenser size. From this standpoint, the size should be determined solely by steady-state requirements.

Figs. 16 through 19 show the effect of varying the magnitude of the load shock with various amounts of refining load, for a condenser size defined by X_f/X_d' = 1.35. Three different magnitudes of load shocks were applied beginning initially with the melting furnace operating at a power factor of 0.707 with respect to the flicker bus. The equivalent resistance of the melting furnace was reduced by (1) 100%, (2) 66²/₃%, and (3) 33¹/₃%, and then returned to its initial value in each case at the most unfavorable time. With respect to swinging of the condenser these three cases are approximately equivalent to a 3-phase, a 2-phase, and a single-phase short circuit respectively of the furnace electrodes, starting from the maximum possible power condition. The corresponding changes of power depend somewhat on all of the system parameters but are approximately as shown in Table I below.

Table I. Effect of Change in Resistance on Power Level in Furnace

% Change of Resistance	% Change of Power
100 (100 to 0)	100 (100 to 0)
66.7 (100 to 33.3)	50 (100 to 50)
33.3 (100 to 66.7)	20 (100 to 80)

The actual magnitudes of the voltage flicker for the partial changes of resistance are not very significant as they are only approximately the average of the flickers on all three phases for the corresponding unbalanced cases.

It is evident that the optimum buffer reactance depends both on the magni-

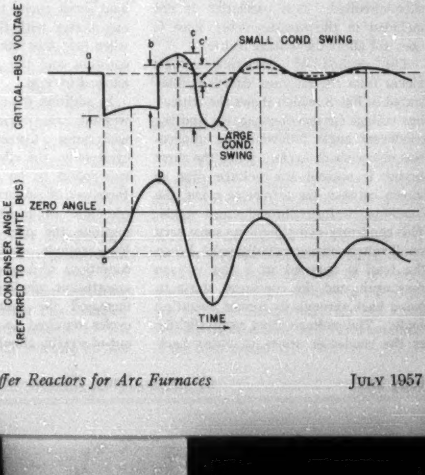


Fig. 8. Voltage-flicker measurement

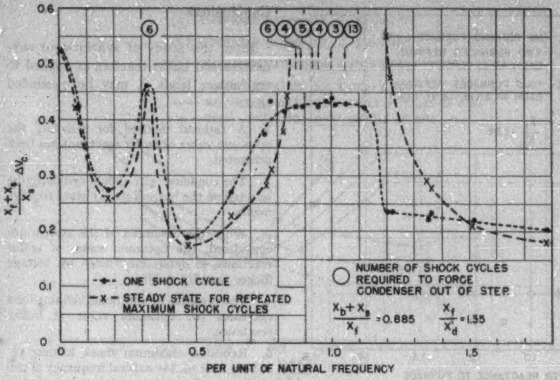


Fig. 9. **Effect of cycling frequency (with maximum shock load) on voltage flicker and system stability**

tude of the load shock and on the amount of additional refining load. This is shown directly in Fig. 20, which summarizes the optimum reactance ratios found from Figs. 16–19. It is further evident that all of the optimum values lie below the value of unity which was suggested as the limit providing an adequate stability margin.

If a single-electrode shock cycle is taken as the typical load shock for which the buffer reactor should be optimized, then for one additional refining furnace the optimum reactance ratio from Fig. 20 is 0.86. This is very close to the value (0.885) in the installation[6] of Fig. 4. Further, from Fig. 17 it is seen that the flicker for the open- to 3-phase short-circuit case has been reduced to one third of the value with no correction. If a maximum 3-phase load shock cycle should occur, Fig. 17 shows that the flicker would be three fourths of the uncorrected open- to 3-phase short-circuit value. Since with no correction this maximum shock would produce a flicker of only half the open- to short-circuit value, this actually represents an increase in flicker over the uncorrected case.

If a maximum shock cycle is taken as the load shock for which the buffer reactor should be optimized, Fig. 20 gives an optimum reactance ratio of 0.55. Then from Fig. 17 for the open- to 3-phase short-circuit case the flicker is only reduced to about 0.44 times the value with no correction, and for the maximum load shock it is reduced from 0.5 to 0.19. However, it does not seem reasonable to use the maximum 3-phase shock cycle as the criterion, since it would occur only very rarely. Moreover, if the maximum load shock and the open- to 3-phase short circuit had been given equal weight, the optimum reactance ratio would be about 0.67 (Fig. 17), which lies between the two previous values.

It is thus seen that the data presented in this paper provide useful information regardless of the load shock used as the basis for selection of the buffer reactance. The method of application is further illustrated in Appendix II.

Finally, it should be remarked that in Figs. 16–20, it is not enough to specify simply the *change* of resistance, or the *change* of power. Instead it is necessary to specify both limits of resistance or power, as is done in Table I. This may be seen not only from the data, but also qualitatively from a very simple steady-state analysis. In the steady state if the condenser maintains its voltages, the only reactive kilovars supplied by the system are half the active kilovars of the buffer reactor. Since the system (or buffer reactor) current (i) is practically proportional to the furnace real power (P), the system reactive kilovars ($\frac{1}{2} i^2 x$) is proportional to the square of this power. The slope of the reactive kilovar versus furnace power curve therefore increases with increasing load, so that a given change in load will produce more change in system reactive kilovars at high powers than at low powers. This is true regardless of the value of buffer reactance, but the effect is only noticeable at high values of buffer reactance. As a corollary to this, the reactive kilovar requirement of the buffer reactor may itself form another limit on its maximum permissible value.

Suggested Future Studies

This paper has not attempted to cover all aspects of arc-furnace application. Some fruitful areas for further investigation are

1. Determination of permissible voltage flicker. This involves the study of the effects of such factors as voltage-time relations and frequency of occurrence of voltage flicker on various sizes and types of lights, on radio and television, and on human reactions to these devices. Suitable definitions, tolerance limits, and measuring instruments need to be developed.

2. Determination of typical real and

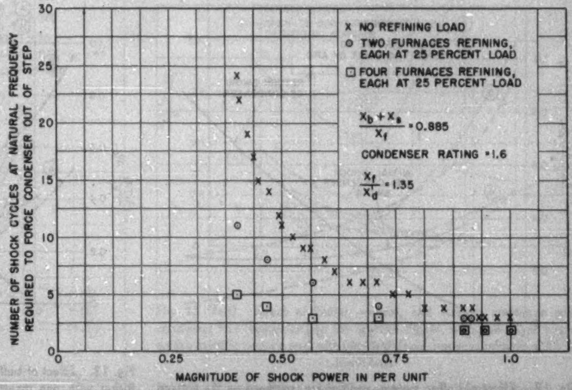


Fig. 10. **Influence of shock magnitude and refining furnace load on stability of system. Condenser rating 1.6 times melting transformer rating. Shocks applied at most unfavorable frequency**

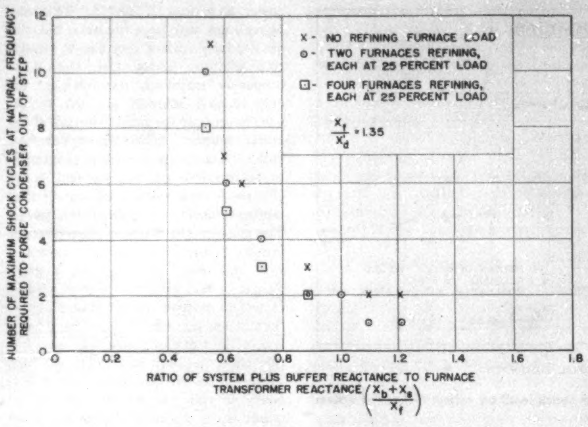


Fig. 11. Influence of buffer, system, and furnace reactance on system stability. Maximum shock applied at most unfavorable frequency

reactive load shock. This involves the study of the relative frequency of occurrence of shocks of various magnitudes and the relations between the real and reactive components of the shocks as affected by furnace reactance, furnace charge, operating technique, and electrode control. There is at present no general acceptance of what should be considered reasonable values of either voltage flicker or load shock. Because of this, the voltage flickers shown in this paper should be interpreted primarily as relative to the voltage flicker obtained with no correction, and not as absolute values.

3. Extension of the present results. The range of system parameters considered can be extended to include larger amounts of steady load. The shock load can be varied to represent more exactly the unbalanced cases and to include other kind of shocks. For example, shocks such as were studied in reference 5 (in which both real and reactive load increased simultaneously) can be evaluated as affected by additional steady loads. For cases which are far outside of the range considered, individual analog computer studies can be made and can be correlated with the present data.

Conclusions

From this study of synchronous condensers and buffer reactors as applied to arc-furnace loads, it may be concluded that:

1. A rational method for selecting the optimum value of buffer reactance has been presented.
2. The condenser size should ordinarily be selected on the basis of steady-state requirements.
3. When the dynamics of the system are considered, the optimum value of buffer reactance is determined more by voltage flicker than by stability.
4. The presence of additional refining load decreases the optimum value of buffer reactance.
5. Repeated maximum shock loading at or very near to, the natural frequency of the system can force the condenser out of step over most of the practical range of system parameters.

Appendix I

Nomenclature

X_s = reactance between the critical bus and the infinite bus
X_b = buffer reactance between the critical bus and the common bus of condenser and furnace transformer
X_f = reactance of melting furnace transformer and leads
X_{fr} = equivalent reactance of refining furnace transformers and leads
R_f = equivalent resistance of melting furnace arc
R_{fr} = equivalent resistance of the refining furnace arcs

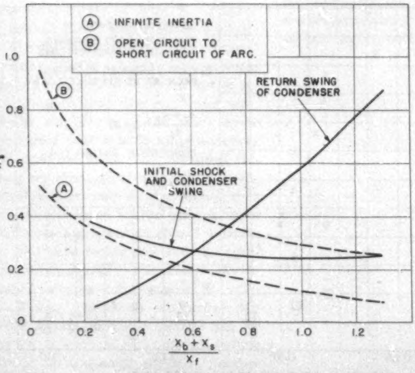


Fig. 12. Effect of buffer, system, and furnace reactance on the voltage flicker, with one maximum shock cycle applied. No refining furnace load, for the case where $X_f/X_d'=1.35$

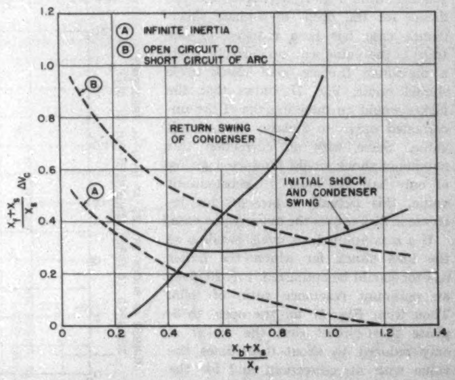


Fig. 13. Effect of buffer, system, and furnace reactance on the voltage flicker with one maximum shock cycle applied on melting furnace. Two additional furnaces refining at 25% load each for the case where $X_f/X_d'=1.35$

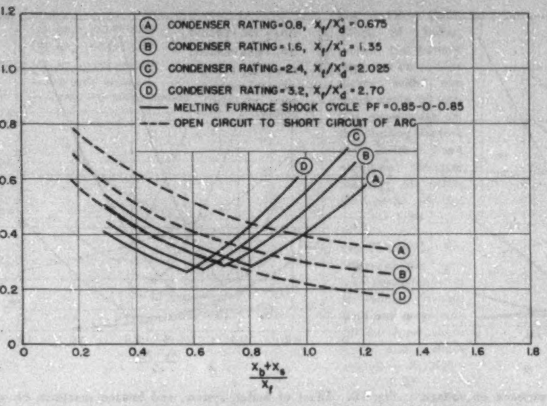


Fig. 14. Effect of buffer, system, and furnace reactance and condenser size (reactance) on voltage flicker, with one shock cycle applied. One furnace melting only

e_{bd}, e_{bq} = direct- and quadrature-axis voltages respectively of the infinite bus referred to the condenser rotor
e_t = terminal voltage of condenser
e_d, e_q = direct- and quadrature-axis components of terminal voltage
i_d, i_q = direct- and quadrature-axis components of condenser armature current
i_{d2}, i_{q2} = direct- and quadrature-axis components of current in the refining furnaces
δ = electrical angle of the condenser with respect to the infinite bus
T_e = synchronous condenser electrical torque
ψ_{df} = field flux linkage
ψ_{kd}, ψ_{nq} = direct- and quadrature-axis amortisseur flux linkages
ψ_d, ψ_q = direct- and quadrature-axis armature flux linkages
ψ_{md}, ψ_{mq} = direct- and quadrature-axis mutual flux linkages
x_d, x_q = direct- and quadrature-axis synchronous reactances
x_{kd}, x_{kq} = direct- and quadrature-axis amortisseur reactances
x_f = field-leakage reactance
x_1 = armature-leakage reactance
r_a = armature-circuit resistance
r_{kd}, r_{kq} = direct- and quadrature-axis amortisseur resistances
r_f = field resistance
E_l = exciter voltage $\left(e_{fd}=\frac{r_f}{x_{ad}}E_l\right)$
E_x = exciter field voltage
M = synchronous condenser inertia
ω_0 = base frequency (377 radians per second)
V_c = component of voltage across the buffer-plus-system reactance in phase with the infinite bus ($V_c=1.0-e_q\cos\delta-e_d\sin\delta$)
ΔV_c = change in critical bus voltage
ΔV_{co} = change in critical bus voltage in going from open- to 3-phase short circuit of furnace electrodes
x_{ad}, x_{aq} = direct- and quadrature-axis mutual reactances
x_d' = direct-axis transient reactance
x_d'' = direct-axis subtransient reactance
T_{do} = field open-circuit time constant
μ = regulator gain
T_a = regulator time lag
T_b = regulator feedback time constant
T_f = exciter field time constant
t = time

System Equations

The system equations were set up on the electronic differential analyzer as follows:

Machine:

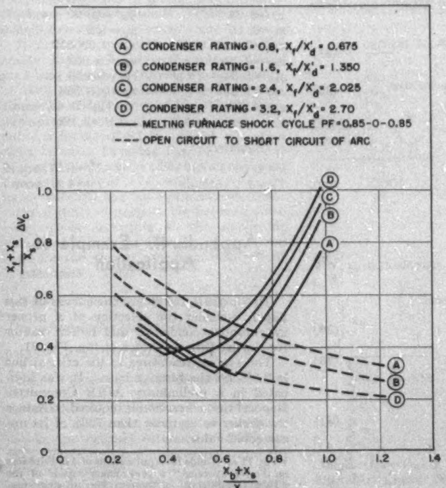
$$\psi_{fd}=\omega_0\int\left[e_{fd}-\frac{r_f}{x_f}(\psi_{fd}-\psi_{md})\right]dt \qquad (4)$$

$$\psi_{kd}=-\omega_0\frac{r_{kd}}{x_{kd}}\int(\psi_{kd}-\psi_{md})dt \qquad (5)$$

$$\psi_d=\omega_0\int\left[e_d+\frac{r_a}{x_1}(\psi_{md}-\psi_d)+\psi_q\right]dt \qquad (6)^*$$

---

* Change in condenser speed was neglected due to its relatively small influence.



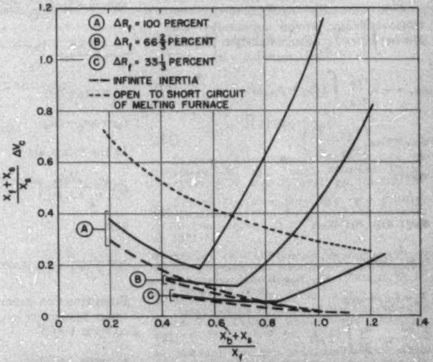


Fig. 15 (left). Effect of buffer, system, and furnace reactance and condenser size (reactance) on the voltage flicker, with one shock cycle on the melting furnace and with two additional furnaces refining at 25% load each

Fig. 16 (above). Effect of buffer, system, and furnace reactance on voltage flicker with one shock cycle of various magnitudes. No refining furnace load. $X_f/X_d'=1.35$

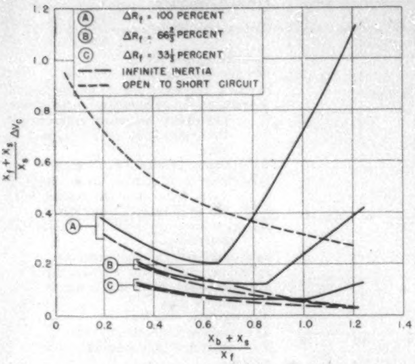


Fig. 17. Effect of buffer, system, and furnace reactance on voltage flicker with one shock cycle of various magnitudes on the melting furnace and with one additional furnace refining at 25% load. $X_t/X_d'=1.35$

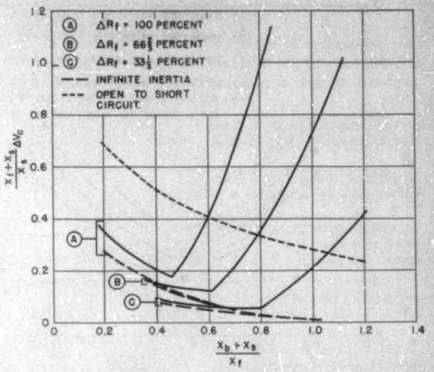


Fig. 18. Effect of buffer, system, and furnace reactance on voltage flicker with one shock cycle of various magnitudes on the melting furnace and with two additional furnaces refining at 25% load. $X_t/X_d'=1.35$

$$\psi_{md}=x_{md}\left(\frac{\psi_{fd}}{x_f}+\frac{\psi_{kd}}{x_{kd}}+\frac{\psi_d}{x_1}-\frac{\Delta(x_{ad}I)}{x_{ad}}\right) \qquad (7)$$

where

$$\frac{1}{x_{md}}=\frac{1}{x_f}+\frac{1}{x_{kd}}+\frac{1}{x_1}+\frac{1}{x_{ad}} \qquad (8)$$

$$\Delta(x_{ad}I)=f(\psi_{md}) \qquad (9)$$

In the foregoing $\Delta(x_{ad}I)$ is taken as the positive difference between $x_{ad}(I_{fd}+I_{kd}-i_d)$ on the no-load saturation curve and the air-gap line.

$$i_d=\frac{1}{x_1}(\psi_{md}-\psi_d) \qquad (10)$$

$$\psi_q=\omega_0\int\left[e_q+\frac{r_a}{x_1}(\psi_{mq}-\psi_q)-\psi_d\right]dt \qquad (11)^*$$

$$\psi_{kq}=-\omega_0\frac{r_{kq}}{x_{kq}}\int(\psi_{kq}-\psi_{mq})dt \qquad (12)$$

$$\psi_{mq}=x_{mq}\left(\frac{\psi_{kq}}{x_{kq}}+\frac{\psi_q}{x_1}\right) \qquad (13)$$

where

$$\frac{1}{x_{mq}}=\frac{1}{x_{kq}}+\frac{1}{x_1}+\frac{1}{x_{aq}} \qquad (14)$$

$$i_q=\frac{1}{x_1}(\psi_{mq}-\psi_q) \qquad (15)$$

$$T_e=i_qe_q+i_de_d \qquad (16)$$

$$\frac{1}{\omega_0}\frac{d\delta}{dt}=-\frac{\omega_0}{M}\int T_edt \qquad (17)$$

$$\delta=\omega_0\int\frac{1}{\omega_0}\frac{d\delta}{dt}dt \qquad (18)$$

Armature circuit:

$$e_q=K_1i_{a1}+K_2i_{b1} \qquad (19)$$

$$e_d=K_2i_{a1}-K_1i_{b1} \qquad (20)$$

where

$$K_1=X_f\frac{\frac{R_f}{X_b}\frac{R_f}{X_f}+\frac{X_f}{X_b}+1}{\left(\frac{R_f}{X_b}\right)^2+\left(1+\frac{X_f}{X_b}\right)^2} \qquad (21)$$

$$K_2=X_f\frac{\frac{R_f}{X_f}}{\left(\frac{R_f}{X_b}\right)^2+\left(1+\frac{X_f}{X_b}\right)^2} \qquad (22)$$

$$i_{a1}=i_d-i_{d2}+\frac{e_{bq}}{X_b} \qquad (23)$$

$$i_{b1}=i_q-i_{q2}-\frac{e_{bd}}{X_b} \qquad (24)$$

$$i_{d2}=c_2e_q+c_1e_d \qquad (25)$$

$$i_{q2}=c_1e_q-c_2e_d \qquad (26)$$

In the foregoing

$$c_1=\frac{R_{fr}}{R_{fr}^2+X_{fr}^2} \qquad (27)$$

and

$$c_2=\frac{X_{fr}}{R_{fr}^2+X_{fr}^2} \qquad (28)$$

From function generators,

$$e_{bd}=\sin\delta=f(\delta) \qquad (29)$$

$$e_{bq}=\cos\delta=f(\delta) \qquad (30)$$

Regulator:

$$e_t=\sqrt{e_q^2+e_d^2} \qquad (31)$$

$$E_x=\mu\frac{T_a}{T_b}(1.0-e_t)+\frac{1}{T_b}\int[\mu(1-e_t)-E_x]dt \qquad (32)$$

$$E_I=\frac{1}{T_f}\int E_xdt \qquad (33)$$

$$e_{fd}=\frac{r_f}{x_{ad}}E_I \qquad (34)$$

Typical System Data Used for One Case

Furnace Transformer: 12,500 kva. (Leads plus transformer reactance = 0.388 on 12.5-megavolt-ampere base on top secondary tap of 285 volts.) On 20-megavolt-ampere base, $X_f=0.621$, $X_s=0.025$.

Synchronous Condenser: 20,000 kva, 13,800 volts.

$x_f=0.38$
$x_1=0.15$
$x_{kd}=0.12$
$x_{kq}=0.13$
$x_{ad}=1.64$
$x_{aq}=1.05$
$x_d=1.79$
$x_d'=0.46$

$x_d''=0.25$
$x_q=1.2$
$r_f=0.000637$
$r_{kd}=0.062$
$r_{kq}=0.056$
$r_a=0.00584$
$T_{do}=8.4$ seconds
$M=1,128$ radians

Regulator:

$\mu=4.56$
$T_a=3.0$ seconds

$T_b=0.5$ second
$T_f=1.0$ second

Appendix II. Example of Application

To illustrate the use of the curves of this paper, consider the selection of a proper synchronous condenser and buffer reactor for the arc furnace shown at top, Fig. 21.

The system reactance to the critical bus is 2.5% on the furnace base. It was indicated in a preliminary AIEE Committee Report† that correction is required to reduce the flicker to no more than 70% of its uncorrected value.

---

* Change in condenser speed was neglected due to its relatively small influence.

† Fig. 11 of "Guide for Applications of Arc Furnaces on Power Systems," a preliminary report of the AIEE Committee on System Engineering, presented at the 1955 AIEE Fall General Meeting.

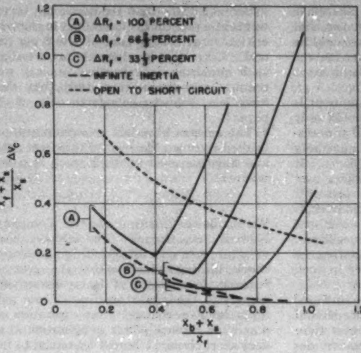
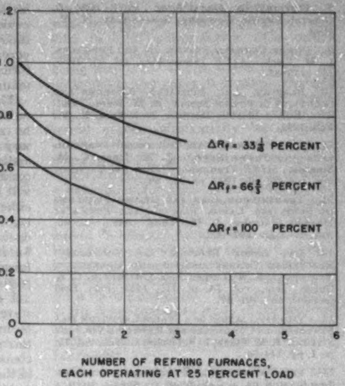


Fig. 19 (left). Effect of buffer, system, and furnace reactance on voltage flicker with one shock cycle of various magnitudes on the melting furnace and with three additional furnaces refining at 25% load each. X_f/X_d' = 1.35

Fig. 20 (right). Relation of the optimum reactance ratios to refining furnace load and magnitude of the load shock (for one shock cycle). X_f/X_d' = 1.35



First the synchronous condenser size will be determined. If the furnace reactance is 50% on its own base, then when it is operating at 120% of rating during melt-down,

$$\sqrt{R_f^2+0.5^2}=\frac{1}{1.2}=0.833$$

or

$$R_f=0.665$$

Therefore the average power factor is 0.665/0.833 or 0.797. The furnace load is 30,000 kva, divided into 23,900 kw and 18,100 kilovars. It is estimated that half the buffer-plus-system kilovars is 5,000 kilovars. Thus the total kilovars to be supplied by the condenser is 23,100. A condenser rating of 25,000 kva is therefore selected, and a condenser transient reactance of 40% assumed. Then $x_f/x_d'=0.50/0.40=1.25$.

It is decided next to select the buffer reactor so as to minimize the voltage flicker for a load shock cycle having a power change of 50% of the maximum possible value. From the middle curve of Fig. 20, the optimum ratio $(x_b+x_s)/x_f=0.85$. Since this curve is for $x_f/x_d'=1.35$ a slight correction must be made by using Fig. 14. Curve *B* of Fig. 14 (for $x_f/x_d'=1.35$) has an optimum reactance ratio of 0.71, while curve *A* (for $x_f/x_d'=0.675$) has an optimum reactance ratio of 0.81. Interpolating between these two curves according to the applicable values of x_f/x_d', an increase of reactance ratio for the present case from 0.85 to 0.87 is estimated. Then

$$x_b+x_s=0.87(x_f)=0.87(0.5)=0.435$$

The actual buffer reactor is

$$0.435-0.025-0.07=0.34$$

(total) − (system) − (transformer) = (buffer)

Recalculation of the kilovolt-amperes consumed in the buffer and system reactance checks the original estimate of 5,000 kva very closely so it is not necessary to modify the condenser size.

Next, the improvement in performance is estimated by means of equation 2.

$$\frac{x_f+x_s}{x_s}\Delta V_{co}=\frac{\left(\frac{0.525}{0.50}\right)}{1+0.87(1+1.25)}=0.355$$

This means that the voltage flicker for an open circuit to 3-phase short circuit is reduced to 35.5% of the uncorrected value. The uncorrected value is found from equation 1 to be

$$\Delta V_{co}=\frac{0.025}{0.525}=0.0477=4.77\%$$

so the corrected value is

$$\Delta V_{co}=0.355(4.77\%)=1.69\%$$

This figure may be checked very easily by considering directly the circuit of the lower diagram of Fig. 21. Equation 2 is merely an evaluation of such a circuit.

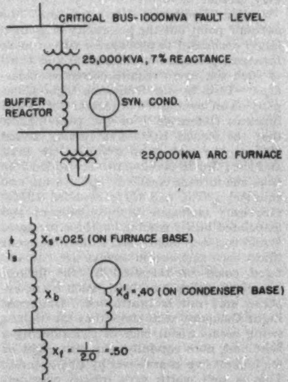


Fig. 21. Circuit diagrams of arc furnace, synchronous condenser, and buffer reactor

The voltage flicker for the load shock used as a criterion may be found from the middle solid curve of Fig. 16, and is

$$\frac{x_f+x_s}{x_s}\Delta V_c=0.125$$

or

$$\Delta V_c=0.125(4.77\%)=0.60\%$$

This is the average of three phases. It is shown in the appendix to reference 5 that for a 2-electrode short circuit the voltage flickers of the three phases have the relative values 2, 0.5, and 0.5, so they would in the present case be 1.20%, 0.30%, and 0.30%.

This example has not taken into account any future expansion, which would require an increased condenser size.

References

1. Electric Arc Furnace Flicker Can Be Overcome, **B. M. Jones.** *Electrical World*, New York, N. Y., Apr. 23, 1951, pp. 114–116, 189.

2. System Capacity Required for 3-Phase Arc Furnaces, **Oliver Rameaur, J. E. Treweek.** *AIEE Transactions*, vol. 73, pt. III-B, Dec. 1954, pp. 1373–74.

3. Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker, **AIEE Subcommittee Report.** *Ibid.*, vol. 76, pt. II, 1957 (*Paper no. 57–9*).

4. Arc-Furnace Loads on Long Transmission Lines, **T. G. LeClair.** *Ibid.* (*Electrical Engineering*), vol. 59, Apr. 1940, pp. 234–42.

5. Voltage Dip and Synchronous-Condenser Swings Caused by Arc-Furnace Loads, **C. Concordia.** *AIEE Transactions*, vol. 74, pt. III, Oct. 1955, pp. 951–56.

6. Arc Furnace Corrective Equipment Using a High Value of Buffer Reactance, **H. W. Harper, T. R. Macon, A. F. Sedgwick.** *Ibid.*, vol. 76, pt. II, May 1957, pp. 74-80.

7. Electronic Analog Computers—d-c Analog Computers (book), **Granino A. Korn, Theresa M. Korn.** McGraw-Hill Book Company, Inc., New York, N. Y., 1952.

8. Distribution Data Book. *GET 1008 B*, General Electric Company, Schenectady, N. Y., 1948.

9. Power Company Service to Arc Furnaces, **L. W. Clark.** *AIEE Transactions*, vol. 54, 1935, pp. 1173–85.

10. Electric Arc Furnaces—Problems and Solutions in Power Supply, **B. M. Jones.** *Bulletin*, Edison Electric Institute, New York, N. Y., Feb. 1951.

11. Large Electric-Arc Furnaces—Performance and Power Supply, **B. M. Jones, C. M. Stearns.** *AIEE Transactions* (*Electrical Engineering*), vol. 60, July 1941, pp. 763–69.

12. Transmission Line and System Problems in Supplying Large Electric-Arc Furnaces During Wartime, **B. M. Jones.** *Ibid.*, vol. 62, Apr. 1943, pp. 197–202.

13. Two Large Electric Arc Furnaces—Electrical Characteristics and Corrective Equipment, **S. W. Luther, J. D. Ghesquiere, C. E. Quick.** *Ibid.*, vol. 74, pt. III, 1955 (Feb. 1956 section), pp. 1401–07.

14. The Application of a Series Capacitor to a Synchronous Condenser for Reducing Voltage Flicker, **P. M. Black, L. F. Lischer.** *Ibid.*, vol. 70, pt. I, pp. 144–50.

15. Technical Problems Associated with the Application of a Capacitor in Series with a Synchronous Condenser, **R. L. Witzke, E. L. Michelson.** *Ibid.*, pp. 519–25.

16. Synchronous Condensers for Steel Mill Service, **E. I. Pollard.** *Iron and Steel Engineer*, Pittsburgh, Pa., 1955, pp. 483–489.

17. Experimental Results with Hollow Electrodes in Electric Steel Furnaces, **W. E. Schwabe.** *Ibid.*, June 1957, pp. 84–92.

◆

Discussion

Benson M. Jones (Consulting Electrical Engineer, Pittsburgh, Pa.): More and larger furnaces are being installed, and this will continue, and therefore it will be necessary for the power-system engineers to determine the method of supplying such arc furnaces in an economical manner to encourage the use of such furnaces. These large furnaces pose real engineering problems for the power company, the furnace builder, and the manufacturer of the associated electric equipment, and the data in this paper will be very useful on many installation problems.

All of us in this part of the electrical industry must co-operate and solve the increasing problems of supplying these big arc furnaces so that they can be installed and used without undue restrictions. It should be kept in mind that all of those involved must profit or it cannot be a satisfactory arrangement.

Power-system engineers, in determining how to supply electric arc furnaces which are large in relation to the stiffness of the supplying electric system, should first set up the tolerable voltage flicker that can be imposed upon their customers. The most critical flicker on general illumination occurs at about six to eight times per second of a magnitude of approximately 0.5 volt. Greater amounts of flicker can also be equally objectionable at a slower rate of occurrence. The flicker limits of the Duquesne Light Company, with which I formerly was associated, are contained in reference 1 of the paper, and the General Electric "Data Book," reference 7 of the paper, contains data on voltage flicker. Next, the engineers should make the necessary system calculations, and measurements as may be needed, including board studies, to determine if the installation will require any corrective devices such as described in this paper, and their size and cost. Such devices should be very carefully engineered jointly with the furnace builder and the electric equipment manufacturer.

The cost of such corrective equipment is quite high and such equipment should only be recommended when absolutely necessary. The authors' data will materially assist in the solution of such problems. It must be remembered that the furnace user will have to pay for such special apparatus either by purchasing, installing, and operating it, or if the power company owns and operates it, then the furnace user will pay for it via his power costs. Naturally, the furnace user is not at all interested in such payments if they can be avoided.

I most heartily agree with the authors' statement that the buffer reactance should be increased because of its lower cost than the synchronous condenser reactance decreased, and their caution on going too far in the buffer reactance increase is important.

As an old power-system engineer who has been exposed to furnace problems, and a former relay engineer, I strongly advise all power-system engineers working on furnace problems requiring synchronous condenser and buffer reactors to analyze their relay and circuit-breaker clearing times and correlate them with the stability of the condenser. The relaying and circuit breakers should be speeded up if necessary; there will be other system benefits from such faster fault clearing.

At this point, I would like to ask the authors if a shunting breaker across the reactor, related to close and short-circuit out the reactor before the condenser becomes unstable, could be a practical matter. Such a shunting breaker has been used, I believe, on some installations for other reasons. With the shunting breaker closed, the condenser would be "stiffer" for other system faults, and as the reactor is not needed most of the time some real benefit may accrue.

Under "Condenser Characteristics," the authors' point out the possibility of using a larger condenser to provide for two or more furnaces, presumably the second and third of such size as to require corrective measures. Tests by the Duquesne Light Company on an installation of several 10,000-kva furnaces (reference 1 of the paper) show that the second 10,000-kva furnace added 20% to the melt-down swing of the first, and the third 10,000-kva furnace added 7%; thus one furnace equalled 100% swing, two equalled 120%, and three equalled 127%. Necessary increases in the condenser and associated buffer reactor for these increases would cost real money and if the infrequent times such increase in swings are encountered could be tolerated by the flicker-conscious utility customers, such extra expense may not be warranted. Duquesne Light Company tests show that the highest swing occurs about once or twice during a heat and, note carefully, the next eight or ten highest swings are lower by 10% or more. This is well worth very serious consideration. Exceeding the power company's flicker limits by a reasonable amount once or twice during a heat of 4 to 6 hours probably could be tolerated. A rather serious furnace swing that is not covered by this paper is the cyclic swing that is of much lower magnitude than the maximum swing, but is at a rate that is sometimes in the very critical range, in the zone of 6 to 8 per second. Corrective equipment to overcome such problems is expensive in cost and requires real engineering study, but they have been solved (see reference 12 of the paper).

The authors have left no stone unturned in their complete analysis of their test data, and have summed up their work in a fine manner.

W. K. Boice (General Electric Company, New Haven, Conn.): The authors have very properly selected for study patterns of shock loading which cause the greatest tendency for swinging of the condenser with respect to the power system. These patterns involve changes from operation at maximum furnace power to operation at or near short circuit, followed by return to full power operation. With such changes, furnace power decreases when reactive load increases.

The paper calls attention to the fact that a practical limit to buffer reactance may be established by moderate but frequent load changes for which furnace power increases when reactive load increases. Such changes increase the kilovars which must be supplied by the power system for consumption by the reactor, and are most important if the total furnace power-load level is high.

It would be helpful to have the authors comment on methods of calculating the optimum buffer reactance as determined by load changes for which reactive and power load increase together.

W. A. Munson (Westinghouse Electric Corporation, East Pittsburgh, Pa.): Upon reviewing the results presented in Figs. 9 through 20 and the example of application given in Appendix II, it appears as though the actual furnace-load power factor has not been taken into account. From an analysis of Figs. 14 and 16, it definitely appears as though the furnace power factor has an effect as the optimum $(X_b+X_s)/X_f$ ratio increases from 0.66 to 0.71 as the power factor increases from 0.707 to 0.85 and the $[(X_f+X_s)/X_s]\Delta V_c$ value also increases from 0.205 to 0.29 for the same power-factor increase. These increases are noted for the X_f/X_d' ratio of 1.35 for both cases with the load variation, i.e., kilowatt variation, being from maximum to zero and then back to maximum again.

As the power factor of the load appears to cause a shift both in the $(X_b+X_s)/X_f$ ratio and the $[(X_f+X_s)/X_s]\Delta V_c$ value, it looks as though a correction factor for power factor as well as for variation in the X_f/X_d' ratio should be employed.

As in the case of X_f/X_d' variation, the power-factor correction can also be interpolated only now the interpolation should be between curve *B* of Fig. 14 and curve *A* of Fig. 16.

Using the example given in Appendix II, and applying a power-factor correction factor as well as an X_f/X_d' ratio-correction factor, it appears as though the optimum reactance ratio for $(X_b+X_s)/X_f$ should be increased to 0.91 instead of 0.87 when taking into account the X_f/X_d' ratio correction only.

This new value for $(X_b+X_s)/X_f$ of 0.91 results in a buffer reactor of 0.36 instead of 0.34 as given in the paper. As this increased value of buffer reactance still results in a reasonable check of the original estimate of 5,000 kva, it is concluded that modification of condenser size is not warranted.

Using this power-factor corrected value of buffer reactance, the improvement in performance as obtained from equation 2 becomes 0.345 instead of 0.355 and results in a corrected value of V_{co} of 1.646% instead of 1.69%.

Again, if we take into account the power factor, we find that the voltage flicker for load shock from Fig. 16 should be 0.758% for the average of the three phases instead of 0.60%. As shown in the appendix of reference 4 of the paper, the voltage flickers of the three phases for the case in question would be 1.516%, 0.379%, 0.379%.

As the power factor appears to have a definite effect both on the flicker level and buffer reactor size, the use of a power-factor correction factor, interpolated from given information, appears to be in order.

C. Concordia, L. G. Levoy, and C. H. Thomas: We wish to thank all of the discussers for their contributions.

Mr. Jones has very properly emphasized the necessity for co-operation among the electric power company, the furnace builder, and the electrical manufacturer. We would like to add to this list the steel company, which is the user of the equipment and thus pays the ultimate bill.

Mr. Jones has also mentioned that the most critical flicker is that which occurs at about 6 to 8 cycles per second with a magnitude of about 0.5 volt. He has in addition stated that this type of flicker is not covered by the paper. Actually, these so-called "cyclic swings" are covered in the paper as much as is any other type of load swing. That is, it is shown that for sufficiently small magnitudes of load swing the static-circuit calculation gives the correct answer for flicker, since the condenser does not swing significantly. This holds for any small load swing regardless of frequency. In fact, if the frequency were sufficiently high the condenser-buffer-reactor arrangement would be even more effective in limiting them, since then the subtransient reactance of the condenser would begin to have an effect on the magnitude of flicker.

We should like to point out that the paper is not primarily concerned with the actual magnitude of load swing that is taken as typical nor with the voltage flicker that is taken as permissible. However, it is necessary to specify both of these factors to arrive at a criterion for satisfactory operation. The magnitude of voltage flicker specified has no significance unless we state the load swing that is used to produce it. Since the load swings produced by arc furnaces are random in nature it is necessary to choose a typical fairly large load swing and specify that with this swing the flicker should not exceed a given magnitude. This indicates that the best guide is experience and we wish to call attention to the AIEE Committee Report, "Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker" (see reference 3 of the paper), as this provides a very good summary of such experience.

In our discussion of this report we have questioned the validity of the shaded areas in Figs. 2, 3, and 4 showing the so-called "cyclic flicker" region for furnaces of rating 15 megavolt-amperes and larger, and will not take space to repeat this discussion here. Note that in the solution of the example in our paper we used Fig. 11 of reference 3 of the paper, which did not show any shaded areas. This report has now been superseded by the foregoing paper in which Fig. 2 without the shaded area duplicates the afore-mentioned Fig. 11.

Mr. Jones has urged the study and possible speeding-up of relay times when condensers and buffer reactors are required. While we agree that such speeding-up will be beneficial to the power system we do not want to leave the impression that the application of arc furnaces on a power system requires the speeding up of relays or circuit breakers.

We do not recommend the use of a shunting circuit breaker across the buffer reactor when the buffer reactor and condenser are put in to alleviate flicker voltages on the system for the following reasons:

1. It is unnecessary if the buffer reactor and synchronous condensers are properly chosen.
2. It would be undesirable to short out the reactor on large swings where its corrective action is most needed.
3. It would put severe stress on the windings of the condenser if it were shorted out when the angular displacement was large.
4. Such action would result in a severe "bump" and voltage dip on the system.

The only time it would be desirable to short out the buffer reactor is if the arc furnaces were not operating, and it was desired to use the condenser for plant or system power-factor correction. Usually such correction is not needed if the furnaces are shut down.

Regarding the condenser size, it is recommended that this be chosen on the basis of steady-state requirements, including present and possible future furnaces with a slight margin to provide for the severe duty. We agree with the general effect of added furnaces on the swing magnitudes as outlined by Mr. Jones, but the condenser must be capable of meeting the steady-state requirements regardless of the way additional furnaces add to the swing magnitude, since it is necessary to maintain the average voltage on the flicker bus at the rated value to make rated production possible.

Mr. Boice has raised a question regarding the calculation of the optimum buffer reactance for load changes involving a simultaneous increase of kilowatts and kilovars to the furnace. To get an increase in kilowatts and kilovars simultaneously represents a change in the impedances of the furnace such as might be obtained by going from 0.85 to 0.707 power factor. The power change here is less than would be obtained in going from 0.707 to zero power factor. As indicated in the paper, it is power changes rather than reactive changes which cause the swinging of the condenser, hence we would expect substantially less angular swing in the condenser than would occur in going from 0.707 power factor to zero power factor, for example. If it is desired to make a calculation for the smaller power shock, the effect of the power shock can be estimated from the curves. To this can be added the effect of the incremental var loading by static circuit theory to get an approximation to the net effect on the critical bus. A more exact method of course would be to set it up on a computer. In fact, reference 4 of the paper contains a study of the response of a very similar system to shocks in which both real and reactive power increase simultaneously.

There is another situation which can give rise to simultaneous increase in kilowatt and kilovar loading as viewed from the critical bus, particularly when high values of buffer reactors are used. A sudden increase in kilowatts of the load served from the flicker bus will necessitate an increase in the power flow through the reactor with the consequent increase in I^2X of the reactor. As a result, viewed from the critical bus a kilowatt increment in the load on the flicker bus results in an increase in both kilowatt and kilovar load as viewed from the critical bus. Here again it is possible to approximate the answers by static circuit theory. Since the shock loading is not known precisely, such approximations are generally satisfactory. Where more exact answers are required for a specified shock-loading pattern, the analog computer can be used.

One additional factor which should be considered in using very large values of buffer reactance is the magnitude of the transient voltage swings on the flicker bus. These swings may become large enough to become objectionable by interfering with the smooth operation of other furnaces operating from the same flicker bus. This effect may in some cases provide an additional limit on buffer reactance.

We can not agree with Mr. Munson's comment and calculation that a correction for power factor should be employed in the example given. While it is true that the average power factor of the furnace given in the example is 0.797, this calculation was made for the purpose of sizing the condenser. However, during the melt the power factor will vary, and the problem states that the calculation is to be made using 50% of the maximum possible value, which corresponds to 0.707 power factor, a figure which the furnace may be operating at momentarily overloaded (not average). Since we are looking for the maximum possible swing we use this value to obtain the correct answer to the problem as stated, and it is inappropriate to employ the power-factor correction proposed by Mr. Munson, for this example.

Diesel Locomotives with Mechydro Hydraulic Transmission for Lightweight Passenger Trains

S. E. STEINBROOK
MEMBER AIEE

FOR MANY YEARS the railroads have been operating their passenger business at a loss, and although they have studied this problem, to date they have not succeeded in reducing this loss appreciably. The locomotive and passenger-car builders have also been considering this problem in an effort to help the railroads turn the loss into a profit. In the last year they have put the results of their study into action by designing and building a new type of passenger car. These cars are lower in initial cost, lighter in weight to reduce the power required to haul them, have a lower center of gravity to permit greater riding comfort and more speed when going around curves, and have been designed to require less maintenance. The cars represent a marked change from designs that preceded them.

To haul these new cars, an existing type of locomotive, with some modifications, could have been used. However, the author's company felt that this was a time to review locomotive design and to develop a locomotive designed for the service.

Passenger locomotives now provide cars with power for traction and steam for heating. Although not all railroads have agreed that head-end power plants for passenger cars are desirable from an operating standpoint, all of the car builders realize that considerable savings in weight and cost are possible if the power plant on each car is replaced by a generator set at the head-end. Once a head-end generator set is installed, the use of electric car heat to replace the steam generator is desirable. From the standpoint of car operation, the only change in locomotive design was the installation of an electric generator set and the elimination of the steam generator.

A review of the diesel engine and electric traction equipment now used showed that very little could be done either to decrease the weight, lower the center of gravity, or reduce maintenance. It was noted that German Railways were successfully using a Maybach hydraulic transmission, which was mounted on the truck and driven by a high-speed diesel engine. The engine and transmission have a record, established in service, of low maintenance costs, low engine fuel consumption, and high availability. The engine is lighter in weight than diesel engines built in the United States. This combination of engine and transmission met all the requirements for which the car builders were designing and it was decided to include them in a locomotive. Baldwin-Lima-Hamilton has a license to use this engine and transmission on locomotives built in this country.

Locomotive Design

Given a clean sheet of paper on which to start, a new, lightweight passenger locomotive has been designed. (Three locomotives have been built, one for the New York Central Railroad and two for the New York, New Haven and Hartford Railroad. While most of the information in this paper is directly applicable to the New York Central locomotive, it also applies to the New Haven locomotives.) There are two diesel engines on this locomotive. The traction engine, rated 1,000 hp (horsepower), is mounted on the truck and delivers all of its power to the rails through a Mechydro hydraulic transmission having four speeds forward and four reverse. The auxiliary engine, rated 570 hp, is mounted in the body of the locomotive and drives a 300-kw a-c generator and the air compressor. The a-c generator provides power for all of the locomotive auxiliaries and for train heat, light, and air conditioning.

Fig. 1 shows this new locomotive alongside a conventional 2,250-hp passenger locomotive. Fig. 2 is a general arrangement drawing of the locomotive. The locomotive is 58 feet, 9 inches long; 10 feet wide; and 11 feet high. It is the same height and has a contour similar to the lightweight "Train *X*" cars it hauls on the New York Central Railroad. The locomotive weighs 193,000 pounds.

The transmission is located beneath the cab floor and is accessible through trapdoors. Directly behind the cab is a compartment in which the traction engine is located. An intercooler radiator fan in the roof blows air down into the compartment and pressurizes it. Access to the engine is obtained through doors in the compartment walls. The auxiliary engine and generator, air compressor, equipment cabinets, and fuel oil pump are

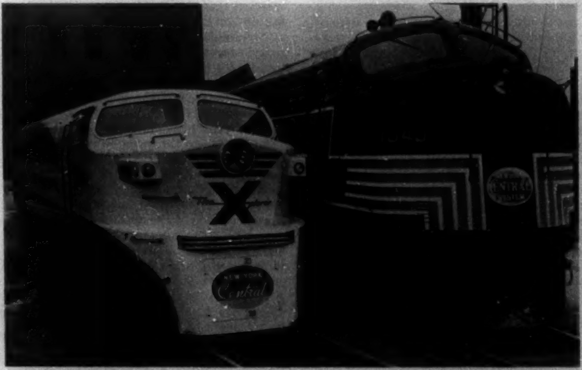


Fig. 1. New lightweight locomotive (left) next to diesel-electric passenger locomotive

---

Paper 57-30, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 26, 1956; made available for printing November 27, 1956.

S. E. STEINBROOK is with the Eddystone Division, Baldwin-Lima-Hamilton Corporation, Philadelphia, Pa.

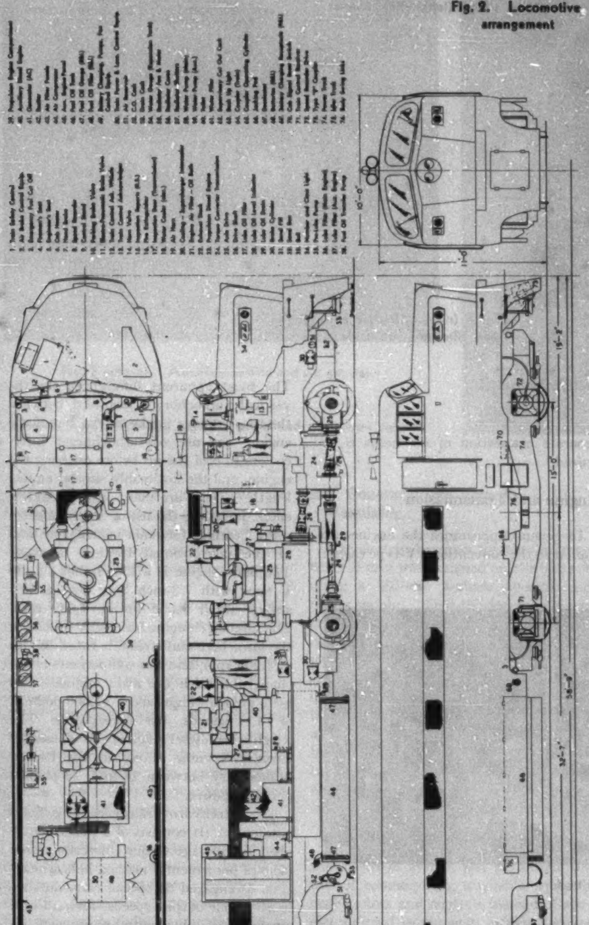


Fig. 2. Locomotive arrangement

located in the center of the locomotive. The fuel oil tank and battery box are below the auxiliary engine and between the trucks. At the rear end are located the fans and radiators for cooling the engines and transmission and the main and intercooler water circulating pumps. The coupler at the front end of the locomotive is a conventional spring-loaded drop type *E*. At the rear end is a special coupler dictated by the design of the cars.

A WABCO-designed type *LF-3-A* automatic air and electric coupler is installed at the rear end of the locomotive (Fig. 3). The coupler and coupler frame on the locomotive are suspended by springs and attached to the unit by four heavy rubber-bushed draw bars, two on each side of the centerline. The coupler contains air brake and 42 75-volt d-c electrical connections, which are automatically connected when the couplers are locked together. No separate connection of wires or hoses is required. Locking of the coupler is accomplished automatically by a pneumatic cylinder mounted on the female portion, which actuates a pawl and engages the male plug. Manual operation of the pneumatic cylinder control valve is required when uncoupling. The three a-c train power contacts are located on top of the coupler frame and mate automatically. In addition to carrying multiple-unit and air-brake lines, the coupler is used to steer the axle of the leading car.

The front truck is a 4-wheel self-contained power package on which are mounted the diesel engine, hydraulic transmission, drive shafts, axle drives, and sand boxes. There is a 4-wheel idler truck at the rear. The power truck is shown in Fig. 4. The locomotive body is suspended from both trucks by body swing links which permit the locomotive to roll into a curve, thus paralleling the roll action of the car. The complete power truck is arranged so that it can be removed and replaced quickly to obtain maximum locomotive utilization.

The driving boxes are $6^1/_2$ by 12 inches on the propulsion truck and $5^1/_2$ by 10 inches on the idler truck, they are equipped with two row roller bearings, grease-lubricated. The top of each box is fitted for the application of a semi-elliptic spring with a resilient pad installed between the spring seat and top of journal box housing. The driving box mounting is of Silent-Bloc arrangement, guided in the pedestal opening by heavy diagonally located rubber-bushed links. One end of the link is attached to the driving box and the other end to the truck frame.

Fig. 3 (left). Rear coupler

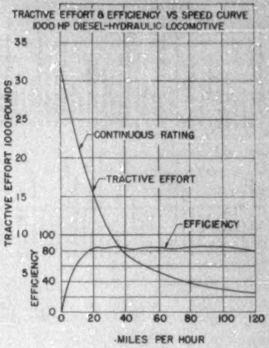


Fig. 6 (right). Tractive effort and efficiency versus speed

The air brakes are electropneumatic, automatic schedule *LWE* with a self-lapping *LWL*-type brake valve for control of the electropneumatic straight air-brake system and the auxiliary automatic emergency brake system. All air-brake lines are contained in the coupler and no separate connection of air hoses is required.

Engine and Transmission

In general appearance the engines are similar to the conventional V-type engine. The most apparent differences are the vertical turbocharger mounted between the two cylinder banks on the top of the engine, the use of turbocharger intercoolers to cool the air before entering the engine, and the noticeably shorter engine length. There are many internal differences, including the use of large-diameter roller bearings for main bearings instead of the conventional friction bearings. The main engine is a 12-cylinder 4-cycle V type with 7.3-inch bore by 7.9-inch stroke, rated 1,000 hp at 1,500 rpm. The auxiliary engine has eight cylinders, the same bore and stroke, is rated 560 hp at 1,200 rpm, and has components interchangeable with the main engine. The main engine weighs about 11,000 pounds, or 11 pounds per horsepower. The weights of diesel engines on passenger locomotives now in use in the United States vary between 16 and 20 pounds per horsepower.

Fig. 4. Power truck

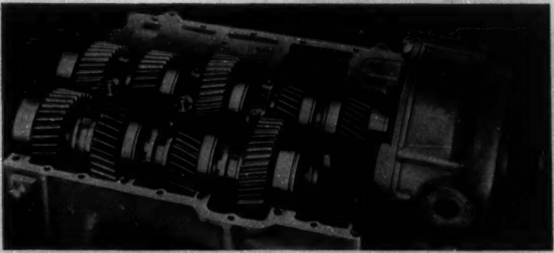


Fig. 5. Transmission gears

The Mechydro transmission is fully automatic. It consists of a permanently filled single-stage torque converter, three pairs of permanently meshed helical gears (Fig. 5) engaged by special claw clutches to give four output speeds, forward and reverse gears, plus control equipment, all housed in a single oiltight casing. Power from the engine to the transmission is through direct-connected universal shafting. The impeller of the torque converter is driven through step-up gearing, thus obtaining a high rotational speed of the converter which results in correspondingly smaller internal parts. Oil used in the transmission is the same as used in the engine. It is cooled by a heat exchanger installed on the transmission which uses water from the main engine circulating system. Power is transmitted to the axle drives from the transmission output flanges by means of heavy-duty longi-

Fig. 7 (above). Auxiliary engine and a-c generator

Fig. 8 (right). Voltage regulator panel and equipment cabinet

tudinal shafts. The axle drives are mounted on the axles in roller bearings. The axle-drive housings are cast-steel designed to protect against external damage and provide maximum support for the gears.

Continuous tractive effort is 21,000 pounds at 12.6 mph. Maximum speed is 120 mph. A higher continuous rating could be obtained at a lower speed by increasing the water cooling rate. Fig. 6 shows the variation in locomotive tractive effort and transmission efficiency with speed. It should be noted that maximum speed is 9.5 times the speed at the continuous rating. On 4-axle diesel-electric locomotives the maximum speed is approximately seven times the continuous speed.

The oil-filled converter permits an immediate response when the throttle is advanced. Locomotive starting is smooth. Operation of the controls in the transmission is fully automatic and is controlled by two governors. One governor is driven by the engine while the other is driven by the output shaft of the transmission and its speed is proportional to locomotive speed. The governors are linked together so that at certain differentials in speeds the claw clutches will operate to make either an upward or downward shift.

The transmission weighs 4,900 pounds. This is considerably less than the weight of a d-c generator and two traction motors. The weight of the generator and traction motors would be about 20,000 pounds.

A-C Generator and Voltage Regulator

The electrical system for the Pullman train *X* cars was designed on the basis of using a 480-volt 3-phase 60-cycle a-c power supply. Since this locomotive has been developed for use with passenger cars only, all of the locomotive auxiliaries, with the exception of the air compressor, are motor-driven. This permits the entire output of the propulsion engine to be available for traction and the use of one a-c generator to supply power for the locomotive and train, and makes it possible to combine the water cooling for both engines and the transmission into one system. This eliminates a water pump, radiator, and cooling fan for each engine.

The standard voltage and frequency of the electrical system has permitted the use of commercially available squirrel-cage motors and controls instead of special railroad equipment, which permits the use of lighter and less expensive equipment. The elimination of all d-c motors, except for fractional-horsepower ones, should keep electrical maintenance low.

The locomotive control circuits are all 75 volts d-c.

Fig. 7 shows the auxiliary engine and a-c generator. The engine and generator are mounted on a common sub-base. This base is supported on the locomotive underframe by flexible rubber mounts to reduce vibration. A flexible coupling is installed between the engine and generator to reduce torsional vibration. The generator is rated at 300 kw, 450 volts, 3 phase, 60 cycles, 0.8 power factor, 1,200 rpm, 10% overload for 2 hours. It has a dripproof enclosure, class *B* insulation, two ball bearings, and is lubricated with Andok *BR* grease. It is 42 inches in diameter and 33 inches long exclusive of shaft extensions. It weighs 4,300 pounds. The d-c exciter is mounted above the generator and is belt-driven. The exciter is rated at 5 kw, 125 volts, 1,750 rpm, and is shunt-wound.

The maximum load on the generator will occur in the winter when all the electric heating will be in use. On a −20-degree-Fahrenheit day the train load will be 270 kw at 1.0 power factor. On such a cold day it is not expected that the radiator fans will be in operation. If one fan is in use the locomotive load will be 55 kw. This load added to the train load is 325 kw. Since the generator is rated at 481 amperes at 0.8 power factor, the load of 325 kw at approximately unity power factor will not overload the generator. On a 110-degree-Fahrenheit summer day when all air conditioning in the cars is in use and when all locomotive fans are in operation, the maximum load on the generator will be about 200 kw.

A static voltage regulator is used. Fig. 8 shows the voltage regulator panel and a portion of one of the equipment cabinets. The direct current in the exciter field is supplied by a self-saturating magnetic amplifier containing a full-wave rectifier. A-c power for the rectifier is obtained from the main generator. The

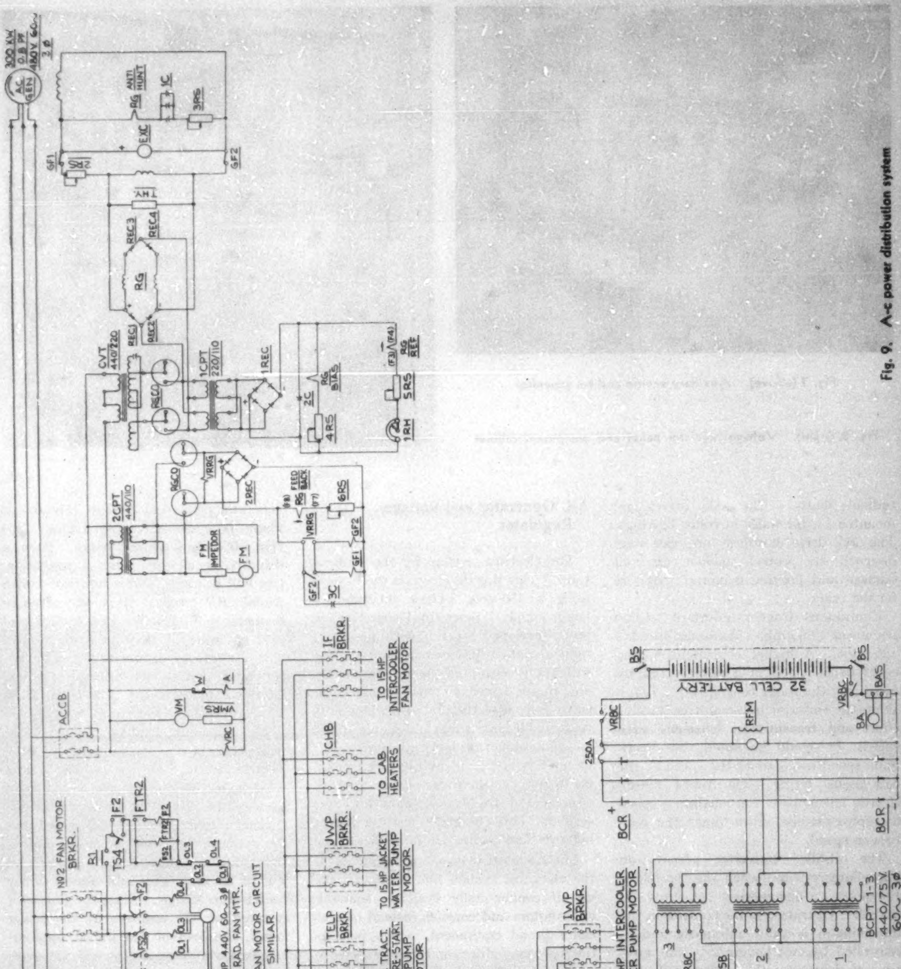

300 KW 0.8 PF 480V 60~ 3Ø
AC GEN
RG ANTI HUNT
1C
3RS
EXC
GF1
VR IN
GF2
1H
REC3
REC4
RG
REC1
REC2
CVT 440/220
RECO
1CPT 220/110
1REC
2C
RG BIAS
4RS
(F3) (F4)
SRS
5RS
RG REF
1RH
2CPT 440/110
RGCO
VRRG
2REC
FEED BACK
RG
6RS
VRRG
GF2
GF1
FM IMPEDOR
FM
3C
ACCB
VM
W
A
VMRS
VRC
1F BRKR.
TO 15 HP. INTERCOOLER FAN MOTOR
CHB
TO CAB HEATERS
JWP BRKR.
TO 15 HP. JACKET WATER PUMP MOTOR
TELP BRKR.
TO 1 HP. TRACT. ENG. PRE-START. LUBE PUMP MOTOR
IWP BRKR.
TO 1 HP INTERCOOLER WATER PUMP MOTOR
NO 2 FAN MOTOR BRKR.
R1
TS4
F2
FTR2
E2
FS2
OL3
OL4
OL2
OL1
40 HP 440V 60~ 3Ø NO 2 RAD. FAN MTR.
NO 1 FAN MOTOR CIRCUIT SIMILAR
MB
TRAINLINE POWER COUPLER
A
W
WPT 1-3 220/4440 V 60~ 3Ø
1
3
2
WAYSIDE RECPT. 220 V 60~ 3Ø
WCB
VRC
A
W
BCCB
VRBC
TESB
BCPT 1-3 440/75 V 60~ 3Ø
BCR +
BCR −
250A
VRBC
RFM
BS
32 CELL BATTERY
BS
VRBC
BAS
BA

Fig. 9. A-c power distribution system

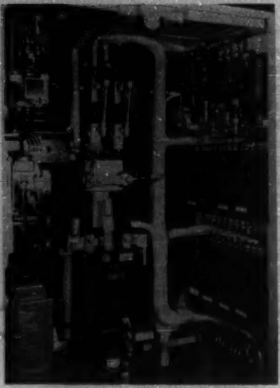


Fig. 10. Circuit breakers and a-c control equipment

output of the magnetic amplifier is controlled by the net ampere-turns supplied by the reference winding of the amplifier (*F3-F4* in Fig. 9), and the feedback winding (*F8-F7*). The voltage across the reference winding (*F3-F4*) is kept constant since it is connected to a constant-voltage transformer. The voltage in the feedback winding (*F8-F7*) varies with generator voltage. As the load on the generator increases the voltage drops and the output of the magnetic amplifier increases, which permits more current to flow to the exciter field. This voltage regulator does not have any moving parts and no maintenance is expected.

A-C Power Distribution System

Fig. 9 shows the a-c power distribution system. Power to the motors is supplied through 3-pole circuit breakers equipped with thermal magnetic trip elements. The train line bus is connected to the generator by means of a motor-operated circuit breaker rated at 400 amperes. It also has thermal and magnetic trip elements. In addition, the main breaker provides undervoltage protection. All of the breakers are calibrated for use in a 50-degree-centigrade ambient. The a-c circuit breakers are shown in Fig. 10.

The two 42-inch fans cooling the engine and transmission water are driven by 40-hp 1,750-rpm squirrel-cage induction motors. The motors have class *B* insulation, a dripproof enclosure, and permanently sealed ball bearings. In order to prevent the generator voltage from dropping excessively when a fan motor

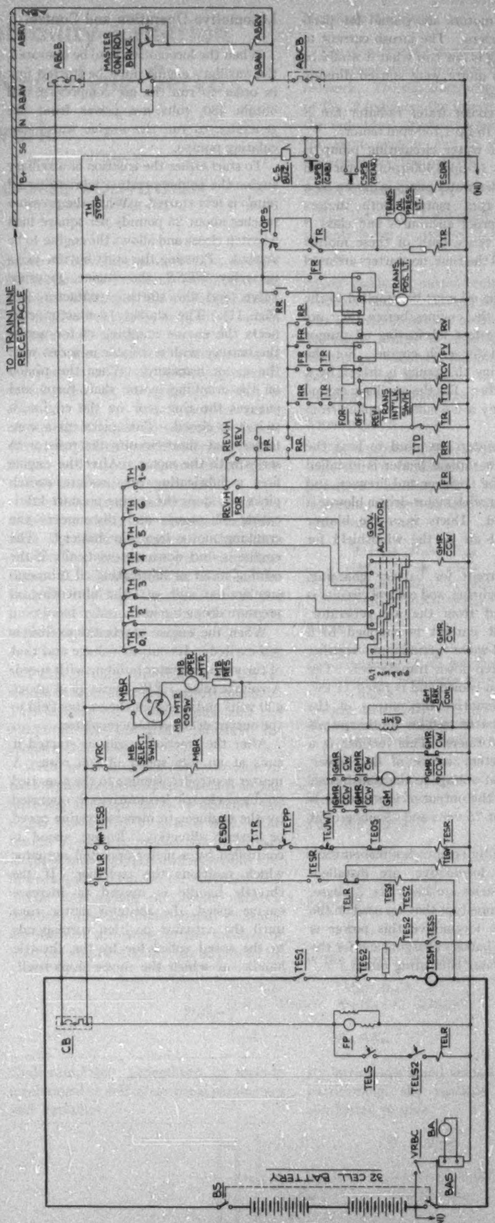


Fig. 11. Locomotive control circuits

starts, the motors are wound for part-winding starters. The inrush current to each winding is one half what it would be if the 40-hp motor were started directly across the line.

The intercooler water radiator fan is driven by a 15-hp 1,750-rpm motor.

The main water circulating pump is driven by a 15-hp 3,400-rpm motor and the intercooler water pump is driven by a 1-hp 3,450-rpm motor. Both motors have splashproof enclosures and class *B* insulation. Since both of these motors are to run all the time, no starters are used with them.

Lubricating oil must be supplied to the bearings of the engines before they are started. Prestart lubricating oil pumps are installed for each engine. For the traction engine this pump is driven by a 1-hp a-c motor. For the auxiliary engine it is driven by a d-c motor supplied from the battery.

Electric heaters are used to heat the cab. A 2-kw radiant heater is installed in front of the engineer and fireman, and a 4-kw heater with motor-driven blower is also installed. Ducts from the blower carry heated air to the windshield for defrosting.

Direct current for battery charging, locomotive lighting, and control circuits is also obtained from the a-c generator. The 480-volt current is supplied to a 3-phase full-wave germanium rectifier through a step-down transformer. The rectifier is fan-cooled and is rated 14 kw. On this locomotive the output of the rectifier is limited to 6 kw by the output of the transformers. This rectifier is a very satisfactory source of d-c power. With an input voltage to the transformer of 480 volts, the output of the rectifier is 15 amperes at 75 volts and 78 amperes at 72.7 volts.

Two stand-by receptacles, one on each side of the locomotive, are installed. These receptacles are 220 volts, 3 phase, and are the same type that are used on the car. On the locomotive this power is only used for battery charging and for the intercooler water circulating pump.

Locomotive Operation and Control

When the locomotive is to be operated, the auxiliary engine must be started first in order to run the air compressor and obtain 480 volts a-c power from the generator to run the engine water circulating pumps.

To start either the traction or auxiliary engine the engine prestart lubricating oil pump is first started. When the pressure reaches about 25 pounds per square inch a switch closes and allows the engine to be started. Pressing the start button picks up relay *TESR*, the engine governor valve, and the starting contactor (see Fig. 11). The starting contactor connects the engine cranking motor across the battery with a resistor in series with the motor armature. When the pinion on the cranking motor shaft turns and engages the ring gear on the engine, a switch is closed. This picks up a contactor that short-circuits the resistor in series with the motor. After the engine fires a lubricating oil pressure switch picks up, stops the engine prestart lubricating oil pump, and disconnects the cranking motor from the battery. The engine is shut down automatically if the cooling water or lubricating oil temperature are too high, or if the lubricating oil pressure drops too low.

When the engine starts the exciter is self-excited. Its output voltage and that of the main generator build up with speed. A voltage relay, *VRRG*, picks up at about 400 volts and transfers the exciter field to the output of the voltage regulator.

After the traction engine is started it runs at an idle speed of 600 rpm. A master controller, similar to the one used on diesel-electric locomotives, is operated by the engineer to increase engine speed or reverse direction. Engine speed is controlled by a motor-operated actuator which positions the governor. If the throttle handle is moved to increase engine speed, the actuator motor runs until the actuator position corresponds to the speed called for by the throttle handle, at which the motor stops itself.

When the locomotive is ready to be moved the reverse lever is moved to forward or reverse. Advancing the throttle from idle to notch *1* energizes a forward or reverse magnet valve on the transmission. This positions the proper directional gear. The transmission control valve is also energized and it permits the converter to engage and transmit torque to the output shaft. Notching out further on the throttle increases engine speed and output. The torque converter disengages the transmission and shifts automatically from one speed to another as determined by locomotive and engine speeds.

The reversing mechanism is protected so that even if the engineer moves the reverse lever from forward to reverse before the locomotive stops, the actual reversal of the gears will only take place at standstill. If the transmission oil temperature rises above 375 degrees, a temperature switch will shut down the traction engine.

Conclusions

A new locomotive designed for passenger service has been built. It provides head-end power for the cars, is lighter in weight, will require less maintenance and reduce operating costs, and uses a hydraulic transmission.

To date, the locomotive has been in revenue service for 137 days with an availability of 92%. It has accumulated 70,055 miles. Fuel consumption for both engines is between 1.0 to and 1.1 gallons per mile.

Sufficient operating experience has been gained both in Europe and the United States to consider the hydraulic-drive locomotive out of the experimental realm. In passenger service where the added advantage of light weight is desirable it appears most appealing to help railroads reduce the costs of operation.

Reference

1. APPLICATION OF THE TORQUE CONVERTER TO LOCOMOTIVES, J. S. Newton. *Mechanical Engineering*, New York, N. Y., June 1954.

Transfer Matrix Stability Criterion Applied to an Amplidyne Servo Network

PIERRE M. HONNELL
MEMBER AIEE

DANIEL WOLFFENSTEIN
STUDENT MEMBER AIEE

Synopsis: The problem of stability rightly enjoys a central role in the theory of servomechanisms. In this paper the transfer matrix stability criterion[1] is applied to an electromechanic servo system employing an amplidyne as the amplifying element. This criterion is valid for either exact neutralization, the strict amplidyne amplifier (in which case a unilateral feedback loop ensues) or for over- or under-neutralization, the Metadyne amplifier (which results in a bilateral transmitting feedback loop). Both cases of neutralization yield substantial agreement between the mathematical predictions and the physical performances, with allowances for the necessary idealizations introduced to render the analysis tractable.

A SERVO NETWORK utilizing a dynamo motor and tachometer with an amplidyne as the sole amplifying element may have a transmission loop with either unilateral or bilateral properties, depending upon the degree of neutralization or compensation of the machine. It is convenient in the first instance, and necessary in the second, to employ the transfer matrix stability criterion if stability of the servo network is to be predicted from a transmission loop analysis. This requires the computation of the transfer matrix $[{}^{0}a]$ of the opened transmission loop as defined by a definite but arbitrarily located cut. The $[{}^{0}a]$ matrix of the open loop yields the quantities entered in the stability criterion.[2]

$$t({}^{\circ}a)-d({}^{\circ}a)-1=0 \tag{1}$$

The roots of equation 1 predict the complex natural frequencies or modes of oscillation of the closed transmission or feedback loop; that is, the servo network. It should be noted that the usual treatment of servomechanisms as a tandem connection of fictitious noninteracting elements is incorrect for a bilateral transmitting feedback loop, as pointed out in the references cited.

To fix ideas and lend concreteness to the analysis, a specific servomechanism is investigated, an ideograph of which is shown in Fig. 1. It is a position servo consisting of input potentiometer transducer, amplidyne amplifier, dynamo servo motor, and output tachometer and position potentiometer transducers.

The case of exact neutralization of the amplidyne yields the simplest analysis, since the unilateralization of the transmission loop introduces some simplifications. It is a simple matter to extend the analysis to include the derivation of the functional relation between the output angle θ_0 and the input angle θ_i, or

$$\theta_0=f(\theta_i) \tag{2}$$

and this permits computation of the response to arbitrary input angle commands. This is done for a step input, and the computed and physically measured responses compared.

With over- or under-neutralization of the amplidyne, strictly the Metadyne, the analysis, although no more difficult, becomes too lengthy to include here. The results of stability predictions for certain of these more complicated situations are, however, also included for completeness.

The appropriate duodinode network matrices and the input-output relations are derived according to the method of Strecker and Feldtkeller,[3] as improved by Guillemin,[4] but generalized to include mechanical as well as electrical parameters and variables.

Network Analysis

The objective of this section is to derive the mathematical relations of the servo network. The simplified schematic shown in Fig. 2 indicates the interconnection of the individual duodinode components of the servo system, and is the point of departure for the analysis.

INPUT TRANSDUCER P_i

For the sake of simplicity, the input potentiometer transducer P_i is defined by

$$u_i=k\theta_i \tag{3}$$

which yields an output voltage u_i proportional to the command input angle θ_i. The internal resistance of this equivalent Thévenin source is neglected since it ultimately becomes insignificant compared to the resistance of the control winding of the amplidyne A.

AMPLIDYNE A

The analysis of the amplidyne starts with Kron's impedance matrix for the primitive machine, Fig. 3(A), and its interconnection as an amplifier, Fig. 3(B). This matrix reads

$$\|{}^{1}Z\|=\left\|\begin{matrix} r_1+L_1\lambda & M_{12}\lambda & M_{13}\lambda & 0\\ M_{12}\lambda & L_2\lambda & M_{23}\lambda & 0\\ M_{13}\lambda & M_{23}\lambda & r_3+L_3\lambda & K_{43}\\ -K_{14} & -K_{24} & -K_{34} & r_4\end{matrix}\right\| \tag{4}$$

where

r_i = resistance of i-th winding
L_i = inductance of ith winding
M_{ik} = mutual inductance between windings i and k
λ = derivative d/dt
K_{ik} = generated voltage in winding k due to current in winding i

To obtain the matrix corresponding to the amplidyne connected as in Fig. 3(B), i.e., as an amplifier, the product

$$\|{}^{2}Z\|=\|{}_{t}C\|\|{}^{1}Z\|\|C\| \tag{5}$$

is formed, where

$$\|C\|=\left\|\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{matrix}\right\| \tag{6}$$

is the connection matrix. This yields

$$\|{}^{2}Z\|=\left\|\begin{array}{cc:c} r_1+L_1\lambda & (M_{12}-M_{13})\lambda & 0\\ (M_{12}-M_{13})\lambda & r_3+(L_2+L_3-2M_{23})\lambda & -K_{43}\\ \hdashline -K_{14} & K_{34}-K_{24} & r_4\end{array}\right\| \tag{7}$$

By letting $u_4=0$ and using the indicated partitioning, the variable i_4 may be eliminated to give

$$\left\|\begin{matrix}u_1\\ u_2\end{matrix}\right\|=\left\|\begin{matrix} r_1+L_1\lambda & (M_{12}-M_{13})\lambda\\ (M_{12}-M_{13})\lambda-\dfrac{K_{14}K_{43}}{r_4} & r_3+(L_2+L_3-2M_{23})\lambda-\dfrac{K_{43}(K_{24}-K_{34})}{r_4}\end{matrix}\right\|\left\|\begin{matrix}i_1\\ i_2\end{matrix}\right\| \tag{8}$$

Paper 57-186, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted July 10, 1956; made available for printing December 10, 1956.

PIERRE M. HONNELL is on the staff of Washington University, St. Louis, Mo., and DANIEL WOLFFENSTEIN is a student there.

This paper summarizes a master's thesis submitted by the junior author to the Sever Institute of Technology of Washington University.

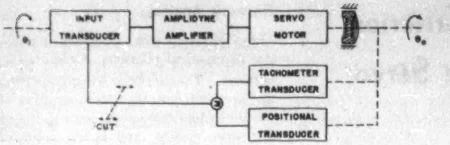


Fig. 1. Ideograph of the servomechanism

Considering first the unilateral case (complete neutralization of axis 3 armature reaction), which means that $L_2+L_3-2M_{23}=0$ and $M_{12}=M_{13}$, it is found that

$$\begin{Vmatrix} u_1 \\ u_2 \end{Vmatrix} = \begin{Vmatrix} r_1+L_1\lambda & 0 \\ -\dfrac{K_{14}K_{43}}{r_4} & r_3-\dfrac{K_{43}(K_{24}-K_{34})}{r_4} \end{Vmatrix} \begin{Vmatrix} i_1 \\ i_2 \end{Vmatrix} \tag{9}$$

From the impedance matrix 9, the transfer matrix for the amplidyne amplifier is found to be

$$[a]=\frac{-r_4}{K_{14}K_{43}}\begin{Vmatrix} r_1+L_1\lambda & (r_1+L_1\lambda)\left\{r_3-\dfrac{K_{43}(K_{24}-K_{34})}{r_4}\right\} \\ 1 & r_3-\dfrac{K_{43}(K_{24}-K_{34})}{r_4} \end{Vmatrix} \tag{10}$$

MOTOR M

The transfer matrix of the driving dynamo with electromechanic co-ordinates as defined in Fig. 4, including all paralleled mechanic parameters (the moment of inertia and friction of the load, gears, and armatures) reads

$$\begin{Vmatrix} v_M \\ i_M \end{Vmatrix} = \begin{Vmatrix} \dfrac{G_M{}^2+\nu_M Z_M}{G_M} & \dfrac{Z_M}{G_M} \\ \dfrac{\nu_M}{G_M} & \dfrac{1}{G_M} \end{Vmatrix} \begin{Vmatrix} \dot{\theta}_M \\ -\tau_M \end{Vmatrix} \tag{11}$$

where

$G_M=P\phi Z/2\pi a$; P = number of poles; ϕ = flux; Z = total number of armature conductors; a = number of parallel paths on armature

$\nu_M=f+J\lambda$; f = coefficient of friction; J = moment of inertia

$Z_M=r_M+L_M\lambda$; electric impedance of armature winding

OUTPUT TRANSDUCERS

The $[h]$ or series-parallel matrix of the tachometer T with co-ordinates defined in Fig. 4 is

$$\begin{Vmatrix} u_T \\ \tau_T \end{Vmatrix} = \begin{Vmatrix} Z_T & G_T \\ -G_T & 0 \end{Vmatrix} \begin{Vmatrix} i_T \\ \dot{\theta}_T \end{Vmatrix} \tag{12}$$

where

$Z_T=r_T+L_T\lambda$, electric impedance of armature circuit

$G_T=P\phi Z/2\pi a$, tachometer generator factor

The idealized $[h]$ matrix of the output potentiometer P_0 (neglecting nonlinear effects) is

$$\begin{Vmatrix} u_p \\ \tau_p \end{Vmatrix} = \begin{Vmatrix} 0 & -\dfrac{U_0}{2\pi\lambda} \\ 0 & 0 \end{Vmatrix} \begin{Vmatrix} i_p \\ \dot{\theta}_p \end{Vmatrix} \tag{13}$$

where U_0 is the (constant) voltage applied to the potentiometer winding.

Addition of the $[h]$ matrices of the tachometer and potentiometer yields the $[h]$ matrix of their combination (the feedback path F)

$$\begin{Vmatrix} u_F \\ \tau_F \end{Vmatrix} = \begin{Vmatrix} Z_T & -\dfrac{U_0}{2\pi\lambda}+G_T \\ -G_T & 0 \end{Vmatrix} \begin{Vmatrix} i_F \\ \dot{\theta}_F \end{Vmatrix} \tag{14}$$

from which the $[b]$ matrix is found to be

$$\begin{Vmatrix} \dot{\theta}_F \\ \tau_F \end{Vmatrix} = \begin{Vmatrix} \dfrac{1}{\dfrac{-U_0}{2\pi\lambda}+G_T} & \dfrac{Z_T}{\dfrac{-U_0}{2\pi\lambda}+G_T} \\ 0 & G_T \end{Vmatrix} \begin{Vmatrix} u_F \\ -i_F \end{Vmatrix} \tag{15}$$

Equations 14 and 15 are each required later in the analysis.

Stability Predictions

When the servo network is cut as indicated in Figs. 1, 2, and 5, the new network thus formed can be considered as a tandem connection of duodinodes for which a transmission or transfer matrix $[^{\circ}a]$ can be defined. The transfer matrix stability criterion applied to this new open transfer loop will predict the modes of oscillation of the closed servo network.

Since the input transducer P_i may be neglected in this formulation, the over-all transfer matrix $[^{\circ}a]$ is obtained by multiplying the individual $[a]$ matrices of the amplidyne A, matrix 10, and motor M, matrix 11; followed by the $[b]$ matrix, matrix 15, of the output transducers which is the $[a]$ matrix of this combined duodinode when viewed from the 2-2′ side. This matrix multiplication of the individual transfer matrices follows the order indicated by the network

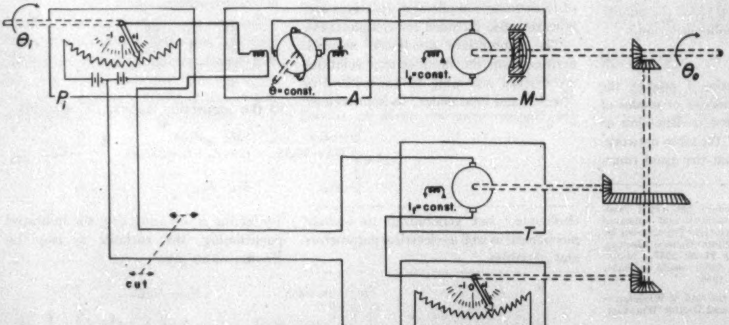


Fig. 2. Schematic of the servo network

graph, Fig. 5, and reads

$$\|{}^\circ a\| = \frac{-r_4}{K_{14}K_{43}} \begin{Vmatrix} r_1+L_1\lambda & (r_1+L_1\lambda)\left\{r_3-\dfrac{K_{43}(K_{33}-K_{34})}{r_4}\right\} \\ 1 & r_3-\dfrac{K_{43}(K_{34}-K_{44})}{r_4} \end{Vmatrix} \begin{Vmatrix} \dfrac{G_M^2+\nu_M Z_M}{G_M} & \dfrac{Z_M}{G_M} \\ \dfrac{\nu_M}{G_M} & \dfrac{1}{G_M} \end{Vmatrix} \begin{Vmatrix} \dfrac{1}{G_T-\dfrac{U_0}{2\pi\lambda}} & \dfrac{Z_T}{G_T-\dfrac{U_0}{2\pi\lambda}} \\ 0 & G_T \end{Vmatrix} \begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix} \tag{16}$$

The negative unit matrix post-multiplier in matrix 16 accounts for the interchange of leads inherent in the topology of the closed transmission loop. The indicated multiplications can be reduced to the following expression

$$\|{}^\circ a\| = \frac{-r_4}{K_{14}K_{43}} \begin{Vmatrix} r_1+L_1\lambda & (r_1+L_1\lambda)[r_3-K_{43}(K_{34}-K_{44})] \\ 1 & r_3-\dfrac{K_{43}(K_{34}-K_{34})}{r_4} \end{Vmatrix} \begin{Vmatrix} \dfrac{1}{G_M} \end{Vmatrix} \begin{Vmatrix} \dfrac{G_M^2+\nu_M Z_M}{\dfrac{U_0}{2\pi\lambda}-G_T} & \dfrac{(G_M^2+\nu_M Z_M)Z_T}{\dfrac{U_0}{2\pi\lambda}-G_T}-G_T Z_M \\ \dfrac{\nu_M}{\dfrac{U_0}{2\pi\lambda}-G_T} & \dfrac{\nu_M Z_T}{\dfrac{U_0}{2\pi\lambda}-G_T}-G_T \end{Vmatrix} \tag{17}$$

which is perhaps as compact as can be written in general terms. At this stage, recourse to numerical coefficients simplifies the analysis.

Upon substitution of numerical values for the parameters of the physical servo network, as listed in the Appendix, with $j=0.1$, equation 17 becomes

$$\|{}^\circ a\| = \frac{-1}{952} \begin{Vmatrix} 89+1.4\lambda & 573+9.02\lambda \\ 1 & 6.44 \end{Vmatrix} \frac{8.5}{15.9-0.256\lambda} \begin{Vmatrix} 0.297\lambda+0.102\lambda^2+0.626\times10^{-3}\lambda^3 & -11.5+3.95\lambda+1.37\lambda^2+3.93\times10^{-3}\lambda^3 \\ 0.1\lambda+0.036\lambda^2 & -4.07+1.4\lambda+0.48\lambda^2 \end{Vmatrix}$$

and finally

$$\|{}^\circ a\| = \frac{-1}{1{,}755-28.67\lambda} \begin{Vmatrix} 83.73\lambda+31.02\lambda^2+0.524\lambda^3 & -3{,}356+1{,}101\lambda+415\lambda^2+6.6\lambda^3 \\ 0.941\lambda+0.334\lambda^2+0.626\times10^{-3}\lambda^3 & -37.7+12.97\lambda+4.46\lambda^2+3.93\times10^{-3}\lambda^3 \end{Vmatrix} \tag{18}$$

where higher order terms are neglected because of their small magnitudes.

The over-all transfer matrix $[{}^\circ a]$, equation 18, yields the coefficients required in the transfer matrix stability criterion,[1] or ${}^0a_{11}+{}^0a_{22}-|{}^0a|-1=0$. In the present instance, the unilateral character of the network simplifies the computations, since the determinant $|{}^\circ a|$ of the transfer matrix vanishes. Thus equation 1 yields

$$\frac{-1}{1{,}755-28.67\lambda}(-37.7+96.7\lambda+35.48\lambda^2+0.528\lambda^3)-1=0$$

and when factored to display the roots, is

$$(\lambda+66)(\lambda+0.6+j7)(\lambda+0.6-j7)=0 \tag{19}$$

The predicted natural frequencies of oscillation therefore are negative real (-66), and conjugate complex $(0.6\pm j7)$. As will be shown later, these results agree with those obtained from the physical network by actual measurement; one exception is that the frequency $\lambda=-66$ gives rise to such a rapidly decaying exponential it cannot be detected in the oscillograms.

Response to Step Input

The function relating the input-quantity θ_i to the output variables can be obtained by adding the $[h]$ matrix of the output transducers to that of the amplidyne and motor combination.

The amplidyne and motor combination $[h]$ matrix can be obtained from the $[a]$ matrix of these components connected in cascade, i.e., the product of matrices 10 and 11. This product reads

$$\|AM_a\| = -\frac{8.5}{952} \begin{Vmatrix} 83.7+31.03\lambda+0.468\lambda^2 & 825+13.14\lambda+2.44\times10^{-3}\lambda^2 \\ 0.941+0.334\lambda & 9.27+1.74\times10^{-3}\lambda \end{Vmatrix} \tag{20}$$

By means of the equivalence

$$\|h\| = \frac{1}{a_{22}} \begin{Vmatrix} a_{12} & |a| \\ -1 & a_{21} \end{Vmatrix}$$

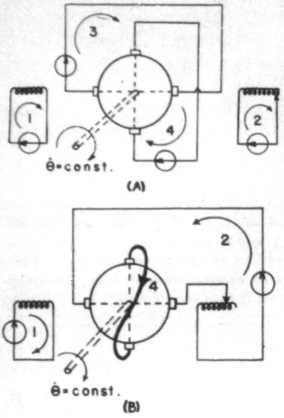
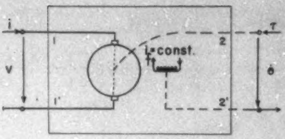


Fig. 3 (left). The amplidyne: (A) primitive; (B) interconnected

Fig. 4 (right). Schematic defining the electromechanic co-ordinates of a dynamo transducer



equation 20 yields

$$\|AM_h\| = \frac{1}{9.27+1.74\times 10^{-3}\lambda}\begin{Vmatrix} 825+13.14\lambda+2.44\times 10^{-3}\lambda^2 & 0 \\ 112 & 0.841+0.33\lambda \end{Vmatrix} \quad 21)$$

the desired $[h]$ matrix of the amplidyne and dynamo motor combination.

Substitution of numerical values into the $[h]$ matrix 14, of the previously derived output transducer combination yields

$$\|F_h\| = \begin{Vmatrix} 13.3+0.0385\lambda & -\dfrac{15.9}{\lambda}+0.256 \\ -0.256 & 0 \end{Vmatrix} \quad (22)$$

Summing matrices 21 and 22 gives

$$\|{}^{\circ}h\| = \frac{1}{(9.27+1.74\times 10^{-3}\lambda)\lambda}\begin{Vmatrix} 948.3\lambda+13.42\lambda^2+2.44\times 10^{-3}\lambda^3 & -147+2.4\lambda+0.445\times 10^{-3}\lambda^2 \\ 109.6\lambda & (0.941+0.334\lambda)\lambda \end{Vmatrix} \quad (23)$$

which relates the output to the input of the complete servo by the scheme given as equation 24

$$\begin{Vmatrix} u_i \\ \tau_0 \end{Vmatrix} = \|{}^{\circ}h\| \begin{Vmatrix} i_i \\ \dot{\theta}_0 \end{Vmatrix} \quad (24)$$

Since $u_i = k\theta_i = (100/2\pi)\theta_i$ the servo output-angle velocity $\dot{\theta}_0$ is given by the following equation

$$\dot{\theta}_0 = -\frac{{}^{\circ}h_{21}}{\|{}^{\circ}h\|}u_i = -\frac{{}^{\circ}h_{21}}{\|{}^{\circ}h\|}k\theta_i \quad (25)$$

in the absence of externally applied torques, or $\tau_0=0$. Upon substitution of numerical magnitudes equation 25 yields the output velocity

$$\dot{\theta}_0 = \frac{-109.6\lambda(9.27+1.74\times 10^{-3}\lambda)\lambda}{(161{,}112+629.4\lambda+329.4\lambda^2+4.52\lambda^3)\lambda}k\theta_i$$

Substituting $k=100/2\pi=15.9$, factoring the denominator, and solving for the output angle

$$\theta_0 = -\frac{109.6(9.27+1.74\times 10^{-3}\lambda)15.9}{4.52(\lambda+71.55)(\lambda^2+1.25\lambda+49.0)}\theta_i \quad (26)$$

which is the over-all servo functional relation, equation 2.

Unit Step Response

The response of the servo network to a unit radian step input angle θ_i is obtained by taking the Laplace transform of equation 26, assuming quiescent initial conditions

$$\theta_0(s) = -\frac{15.9(224.3+0.039s)}{(s+71.55)(s^2+1.25s+49.9)}\theta_i(s) \quad (27)$$

and substituting therein the transformed input $\theta_i(s)=1/s$. Solving for the output angle $\theta_0(s)$ and expanding the result into partial fractions yields equation 28

$$\theta_0(s) = -\frac{1}{s}+\frac{9.67\times 10^{-3}}{s+71.55}+\frac{1.93+0.991s}{s^2+1.25s+49.9} \quad (28)$$

for which the corresponding time function is as follows:

$$\theta_0(t) = -1+9.67\times 10^{-3}\epsilon^{-71.55t}+ 0.991\epsilon^{-0.625t}\{\cos 7.05t+0.188\sin 7.05t\} \quad (29)$$

This computation agrees with results obtained by actual measurement on the physical servo as shown in Fig. 6, a photograph of the dynamic response to a step input. The oscillogram consists of a unit step (the steady-state value), a rapidly decaying exponential term, and a decaying sinusoid, as was predicted by equation 29.

It may also be noted that the natural modes of oscillation computed from the transfer matrix stability criterion correspond to the modes given by equations 28 through 29, allowances being made for the neglect of higher ordered λ-terms in the latter.

Stability of Bilateral Servo Network

Two additional cases were investigated for stability:

1. Overcompensation of the amplidyne yielding threshold of oscillations of the servo network.
2. Undercompensation of the amplidyne yielding an overdamped response.

The method of analysis is identical with that for the case of exact neutralization which gave a unilateral network, except that the transfer matrix of the amplidyne is now necessarily different in each instance. For these bilateral transmitting cases, the amplidyne $[Z]$ matrix reads

$$\|Z\| = \begin{Vmatrix} r_1+L_1\lambda & M_{12}\lambda & M_{13}\lambda & 0 \\ M_{12}\lambda & r_2+L_2\lambda & M_{23} & 0 \\ M_{13}\lambda & M_{23}\lambda & r_3+L_3\lambda & K_{43} \\ -K_{14} & -K_{24} & -K_{34} & r_4+L_4\lambda \end{Vmatrix} \quad (30)$$

which differs from equation 4 in that the terms r_2, r_3, and $L_4\lambda$ are included in the 22, 33, and 44 elements of the matrix 30 thereby slightly increasing the exactness of its representation of the physical device. Upon substitution of measured numerical parameters, this leads to the two amplidyne transfer matrices, as follows:

1. Overcompensated:

$$\|A_a\| = -\frac{1}{952}\begin{Vmatrix} 89+1.6\lambda & 609+4.6\lambda \\ 1+2.2\times 10^{-3}\lambda & 6.85+5.08\times 10^{-3}\lambda \end{Vmatrix} \quad (31)$$

2. Undercompensated:

$$\|A_a\| = -\frac{1}{952}\begin{Vmatrix} 89+1.6\lambda & 3{,}130+4.64\lambda \\ 1+2.2\times 10^{-3}\lambda & 35.25+7.17\times 10^{-3}\lambda \end{Vmatrix} \quad (32)$$

The gear trains used in the physical

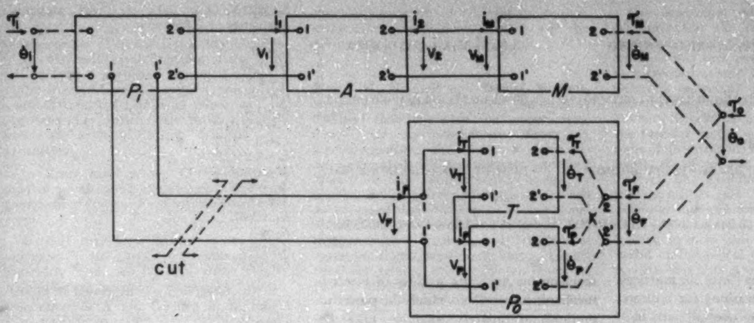


Fig. 5. Schematic defining the electromechanic co-ordinates of the servo network

servo network introduce a highly nonlinear component of friction, and it is not possible to include the effects of these forces by means of a single number. Quantitative agreement between the analysis and the physics can be obtained only by judicious choice of the friction coefficient in each instance from the experimentally determined friction-torque versus angular-velocity curve of Fig. 7.

The appropriate coefficients of friction taken from this curve are respectively $f=0.045$ for the overcompensated, and $f=0.4$ for the undercompensated case.

The over-all transfer matrices of the complete servo networks, neglecting high-order terms, thus become equations 33 and 34, given on the following page.

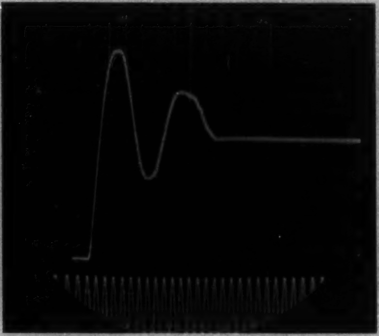


Fig. 6 (left). Servo response to a unit angle step input, exact neutralization; 5-cps time marks

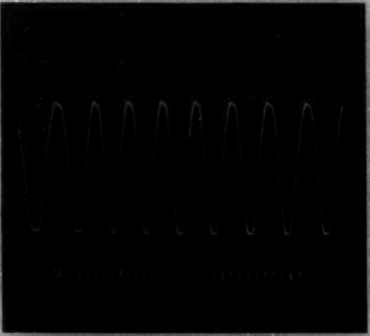


Fig. 8 (right). Servo response to a unit angle step input, overneutralized; 5-cps time marks

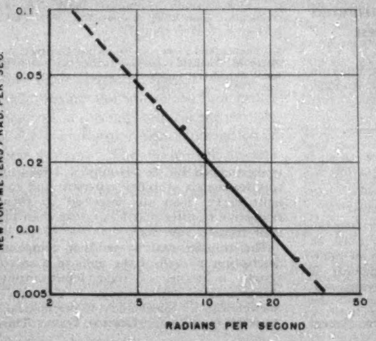


Fig. 7 (left). Experimentally measured friction in network: torque versus angular velocity

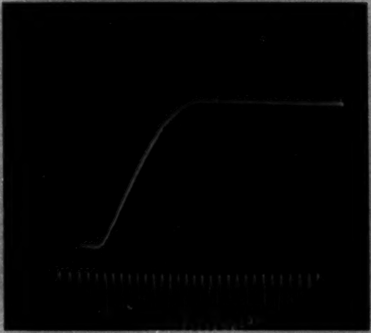


Fig. 9 (right). Servo response to a unit angle step input, underneutralized; 5-cps time marks

1. Overcompensated:

$$\|{}^\circ a\| = \frac{-1}{1{,}755-28.67\lambda}\begin{Vmatrix} 41.5\lambda+31.4\lambda^2+0.385\lambda^3 & -3{,}500+551\lambda+421\lambda^2+6.32\lambda^3 \\ 0.45\lambda+0.348\lambda^2+1.033\times10^{-3}\lambda^3 & -39.42+6.57\lambda+4.68\lambda^2+0.0272\lambda^3 \end{Vmatrix} \quad (33)$$

2. Undercompensated:

$$\|{}^\circ a\| = \frac{-1}{1{,}755-28.67\lambda}\begin{Vmatrix} 1{,}353.6\lambda+125.4\lambda^2+0.386\lambda^3 & -1{,}024+1{,}356\lambda+150.4\lambda^2+2.18\lambda^3 \\ 15.24\lambda+1.38\lambda^2+1.11\times10^{-3}\lambda^3 & -155+205.3\lambda+18.4\lambda^2+0.055\lambda^3 \end{Vmatrix} \quad (34)$$

Since the determinants $|{}^\circ a|$ of the matrices 33 and 34 are nonvanishing for nonzero λ's, it is clear that the over-all (cut) networks have bilateral transmitting properties. That is, these open servo loops will transmit signals in either direction, albeit not equally well.

Application of the transfer matrix stability criterion in these cases yields the following (factored) expressions.

1. Overcompensated:

$$(\lambda+87.6)(\lambda+j6.8)(\lambda-j6.8)=0 \quad (35)$$

2. Undercompensated:

$$(\lambda+318.05)(\lambda+3.12)(\lambda+4.82)=0 \quad (36)$$

These computed results compare favorably with physical measurements. In the first case the sustained oscillations seen in the oscillogram of Fig. 8 correspond to the conjugate imaginary frequencies predicted by equation 35. The oscillographic record in Fig. 9 indicates the response of the closed servo network to a step angular input angle for the second case. In this instance no oscillations but merely decaying exponentials are present, which agrees with the predictions of equation 36 that only real frequencies will occur.

Conclusions

The demand for more exact analyses of servomechanism performance can be satisfied only by correspondingly more rigorous mathematical formulations. This forces the recognition of servomechanisms as being a class of electromechanic networks to which the powerful methods of matrix analysis may be applied. This is exemplified in this paper and in a recent publication describing a linear-motion device.[5] In both instances, theory is correlated with a specific physcial device. While not elaborated upon, it should be clear that the servo equations developed herein carry the formulation of electromechanic equations to a fair degree of sophistication.

The price paid for a more rigorous analysis is the increased complexity of the equations, as should be expected. In this paper manageable results were obtained without recourse to computing machinery by a ruthless excision of higher-order terms. But this was done knowingly for purposes of expediency and not in ignorance and, in principle, it need not have been required.

The experimental results obtained indicate that nonlinear effects, particularly friction, although not vitiating the analysis, do blunt the fine points of comparison between the theory and the physics. Thus these nonlinearities must ultimately be included, and this will undoubtedly be assisted by the matrix formulation of the electromechanic network equations.

Appendix. Machine Parameters as Determined by Test

Amplidyne *A*, a product of the Emerson Electric Manufacturing Company, is model *5AM31NJ18A;* 60 volts, 8.8 amperes, 530 watts at 8,300 rpm.

$r_1=89$
$r_2=0.25$
$r_3=1.0$
$r_4=1.0$
$L_1=1.4$
$L_2=6\times10^{-3}$
$L_3=5.7\times10^{-3}$
$L_4=2.2\times10^{-3}$
$M_{12}=0.0575$
$M_{13}=0.061$
$M_{23}=5.56\times10^{-3}$
$K_{14}=140$
$K_{24}=6$
$K_{34}=6.5$
$K_{43}=6.8$

In the foregoing, r is expressed in ohms, L and M in henrys, and K in volts per ampere.

Servo motor *M*, a product of the Emerson Electric Manufacturing Company is model *305975;* 60 volts, 8.2 amperes, at 4,000 rpm; field 24 volts.

$r_M=2.83$ ohms

$L_M=1.74\times10^{-3}$ henry

$G_M=1/8.5$ newton-meter per ampere

$J_M=0.036$ kilogram-meter^2

Tachometer *T* is a product of the Electric Motor Corporation, model *SPN 44461;* 24 volts.

$r_T=13.3$ ohms

$L_T=0.0385$ henry

$G_T=0.256$ newton-meter per ampere

References

1. The Generalized Transmission Matrix Stability Criterion, **Pierre M. Honnell.** *AIEE Transactions,* vol. 70, pt. I, 1951, pp. 292–98.

2. The Transmission Matrix Stability Criterion, **Pierre M. Honnell.** *The Matrix and Tensor Quarterly,* Kent, England, vol. II, no. 1, Sept. 1951, pp. 8–15.

3. Grundlagen der Theorie des allgemeinen Vierpols, **F. Strecker, R. Feldtkeller.** *Elektrische Nachrichten-Technik,* Berlin, Germany, vol. 6, no. 3, Mar. 1929, pp. 93–112.

4. Communication Networks (book), **E. A. Guillemin.** John Wiley & Sons, Inc., New York, N. Y., vol. 2, 1935, chap. IV.

5. Prescribed-Function Vibration Generator, **Pierre M. Honnell.** *Journal,* British Institution of Radio Engineers, London, England, vol. 16, no. 4, Apr. 1956, pp. 187–98.

◆

Discussion

R. J. W. Koopman (Washington University, St. Louis, Mo.): Servomechanism synthesis procedures built around the complex-plane locus of the open-loop transfer function or similar diagrams and the modification of these diagrams by the addition to the loop of units described by transfer functions are well known.

These methods have two principal limitations:

1. The system must be unilateral.
2. The loading of the units forming the loop by other units in the loop must be negligible, prevented by the use of additional or more expensive equipment, or accounted for by rather annoying procedures.

In this paper, stability in a servo system is determined by Dr. Honnell's 2-way or transfer matrix stability criterion, and the units in the loop are described by their respective transfer matrices rather than by their transfer functions.

The transfer matrix yields a complete description of each of the units of a servo system in terms of their input-output characteristics, whether they be electrical-to-electrical, mechanical-to-mechanical, electrical-to-mechanical, or vice versa. This facilitates calculation of the effects of signals traveling both in the forward and in the reverse sense. The 2-way stability criterion of Dr. Honnell permits cutting of the loop, for analytical purposes, at any point.

These considerations suggest where the methods employed in this paper can be applied and expanded in the future development of servomechanism analysis and synthesis:

1. They may prove to be valuable aids in partitioning the representations of complicated systems for piecemeal analytical study.

2. One may imagine that with future advances in the theory and application of matrices, and as facility with matrix manipulation increases, especially through machine methods, a whole new synthesis art, paralleling that mentioned in the opening paragraph of this discussion but without its limitations, might well be developed.

Pierre M. Honnell and **Daniel Wolffenstein:** Dr. Koopman clearly summarizes the new directions which our paper introduces into servomechanism theory. A servomechanism, by whatever name, is an integrated electromechanic network system, and to be represented accurately analytically requires mechanical as well as electrical co-ordinates. The functional relationship between the dependent and independent electrical and mechanical variables along these co-ordinates is most concisely related by the matrix calculus, thereby also making accessible new results such as the bilateral transfer matrix stability criterion.

The researches described in the paper implement the remarks in the last paragraph of reference 1, wherein the new criterion was first described, as applicable to servomechanisms as well as to purely electronic networks. Furthermore, our former students, Dr. Sefik Altay, and Dr. Norbert Schalk, have applied with success the transfer matrix stability criterion and some of its variations to transistor networks, which are in general bilateral.

Thus, Dr. Koopman's predictions of a new synthesis art for servomechanism networks are well founded: a more exact and more accurately formulated theory leading to more exact and therefore more useful mathematical methods of analyzing servomechanisms doubtless will be along these lines, at a time when electrical engineers need once again to lead the way to new horizons in the field.

Compensation of an Error-Sampled System by a Multirate Controller

GEORGE M. KRANC
NONMEMBER AIEE

Synopsis: A special type of multirate sampled-data system, i.e., a system which contains switches operating with different sampling periods, is discussed in this paper. It is shown, after a brief outline of the multirate mathematics, how a multirate controller can be designed to compensate an error-sampled feedback system. Such a controller can be realized as a pulsed network which in response to input samples at time $t=0, T, 2T, \ldots$ produces output samples at time $t=0, T/n, 2T/n, \ldots$ (where n is an integer), any output sample being a linear combination of the past and present inputs and outputs. The multirate controller achieves a lower intersampling ripple and a faster settling time in comparison with its single-rate predecessor at no increase, in most cases, in the number of the required storage elements.

THE MODERN TREND in the design of error-sampled feedback systems[1-4] tends toward the use of discrete rather than continuous networks for the purpose of the compensation and stabilization of these systems. The use of such discrete networks permits the designer to operate completely in the time domain,[3,4] frees him from the limitations imposed by the stringent physical realizability requirements of continuous networks, and lessens greatly the computational labor. In designing a discrete compensator, one has a large degree of freedom in prescribing the over-all response of the system to specified inputs at the sampling instants. Little or no control can be usually exercised over the values of the continuous output between the sampling instants. Thus, although it is possible to realize, by means of a discrete compensator, a system whose output follows the input at sampling instants in the steady state, this does not preclude the existence of intersampling ripple, which may be undesirable. In order to suppress the ripple, the discrete compensator must provide some output information between the sampling instants, or, in other words, it must behave like a continuous network.

The multirate controller, described in this paper, achieves the necessary compromise between a continuous and a discrete network. Such a controller delivers output pulses between the sampling instants of the uncompensated system and causes the plant to be excited by a smoother signal than in the case of the more conventional single-rate controller.[3,4] In this way the plant output tends to have a lower ripple content.

The presence of the multirate controller implies the use of different sampling rates in the same system. Such systems are classified as multirate sampled systems and it has already been shown that they can be handled mathematically by S-transform methods.[5] A review of the multirate mathematics applicable to the class of systems described in this paper is given in the next section.

Mathematical Background

Assuming as usual that each sampling switch can be replaced by an impulse modulator, it can be written by inspection, from Fig. 1(A), that

$$c(t)=\sum_{k=0}^{\infty}e(kT)h(t-kT) \qquad (1)$$

where $e(t)$ and $c(t)$ are input and output of the network H, respectively, and $h(t)$ is the impulsive response of H. To obtain the sampled output, $c(mT/n)$, t is replaced by mT/n in equation 1. Thus,

$$c(mT/n)=\sum_{k=0}^{\infty}e(kT)h(mT/n-kT) \qquad (2)$$

In analogy to the Z-transform of $c(t)$, $C(z)$, which is sometimes written as $Z[c(t)]$ or $Z[C(s)]$ and is defined as

$$C(z)\triangleq\sum_{m=0}^{\infty}c(mT)z^{-m} \qquad (3)$$

where $z\triangleq e^{sT}$, it is convenient to define the Z_n-transform of $c(t)$ as

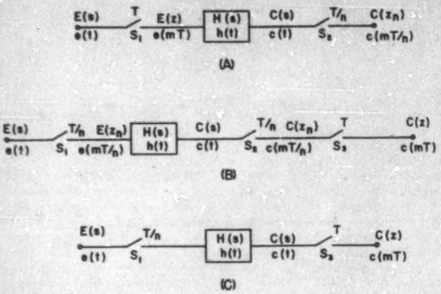
$$C(z_n)\triangleq\sum_{m=0}^{\infty}c(mT/n)z_n{}^{-m} \qquad (4)$$

where $z_n\triangleq e^{s/Tn}=z^{1/n}$. The Z_n-transform

Paper 57-187, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted September 27, 1956; made available for printing January 8, 1957.

George M. Kranc is with Columbia University, New York, N. Y.

The author wishes to thank Professors Ragazzini and Franklin and Mr. Bertram for their useful advice and interest in this research. This work was supported by the United States Air Force under Contract no. *AF 30(635)-2815* and monitored by the Rome Air Development Center, Griffiss Air Force Base, Rome, N. Y.



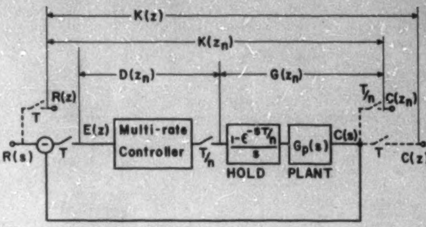


Fig. 1(left). Signals in open-loop multirate systems

Fig. 2 (above). Error-sampled system using multirate controller

of $c(t)$ can also be written as $Z_n[c(t)]$ or $Z_n[C(s)]$. Z_n-transforming equation 2 in accordance with the definition of the transformation given by equation 4, it follows that

$$C(z_n)=\sum_{m=0}^{\infty}\sum_{k=0}^{\infty}e(kT)h(mT/n-kT)z_n^{-m} \qquad (5)$$

Replacing $mT/n-kT$ by pT/n in equation 5, it is possible to rewrite it as

$$C(z_n)=\sum_{k=0}^{\infty}e(kT)z^{-k}\sum_{p=0}^{\infty}h(pT/n)z_n^{-p} \qquad (6)$$

and it follows therefore from the definitions of the Z- and Z_n-transforms that

$$C(z_n)=E(z)H(z_n) \qquad (7)$$

where $E(z)$ and $H(z_n)$ are, respectively, the Z-transform of $e(t)$ and the Z_n-transform of $h(t)$. It is observed that $H(z_n)$ can be obtained in closed form from tables of Z-transforms by replacing all z by z_n and all T by T/n in the expression for $H(z)$, the Z-transform of $h(t)$. This procedure is, however, restricted to those responses $h(t)$ which have no time delay. The procedure to be adopted for responses with delay is outlined in Appendix I.

It follows from the definition of z_n that $z=z_n^n$. Equation 7 can thus be written

$$C(z_n)=E(z_n^n)H(z_n) \qquad (8)$$

where $E(z_n^n)$ is obtained by replacing every z by z_n^n in the expression for $E(z)$. Thus, in general, $C(z_n)$ will appear as a ratio of two polynomials in z_n^{-1} and standard methods of real inversion such as long division, partial fraction expansion, or contour integration can be applied to obtain $c(mT/n)$.

Equation 7 can also be derived by use of the periodic properties of Z-transforms. Thus, it can be directly written from Fig. 1(A) that

$$C(z)=E(z)H(z) \qquad (9)$$

and therefore that

$$C(z_n)=Z_n[C(s)]=Z_n[E(z)H(s)] \qquad (10)$$

But, just as any Z-transform is periodic in the frequency domain with the period $\omega_0=2\pi/T$ so the Z_n-transform is periodic with the period $n\omega_0$. Therefore, since $E(z)$, being periodic with respect to ω_0, must also be periodic with respect to $n\omega_0$, it follows that equation 10 can be rewritten as

$$C(z_n)=E(z)Z_n[H(s)] \qquad (11)$$

which is the same as equation 7.

Considering next the system shown in Fig. 1(B), it can be written that

$$C(z_n)=E(z_n)H(z_n) \qquad (12)$$

Equation 12 follows from the fact that both switches S_1 and S_2 operate with the same sampling period T/n. The output switch S_3 does not perform any sampling operation but rather acts as a selector, letting through only every nth sample out of the sequence $C(z_n)$.

Let $P_n[C(z_n)] \triangleq$ "every nth term of $C(z_n)$." Sklansky[6] has shown that

$$P_n[C(z_n)]\equiv\frac{1}{n}\sum_{k=1}^{\infty}C\left(z_n\epsilon^{\frac{j2\pi k}{n}}\right) \qquad (13)$$

Observing that $P_n[C(z_n)]$ must be identically equal to $C(z)=Z[C(z_n)]$, where $Z[C(z_n)]$ is the Z-transform of a function $c(t)$ such that $C(z_n)=Z_n[c(t)]$, it can be written that

$$C(z)=\frac{1}{n}\sum_{k=1}^{\infty}E\left(z_n\epsilon^{\frac{j2\pi k}{n}}\right)H\left(z_n\epsilon^{\frac{j2\pi k}{n}}\right) \qquad (14)$$

It is obvious from Fig. 1(B) that short-circuiting of the switch S_2 has no effect on the output from the switch S_3 and therefore the expression for $C(z)$ given by equation 14 also applies in the case of the system in Fig. 1(C). The system in Fig. 1(B) can be thought of as a useful mathematical analog of the system in Fig. 1(C).

The identity stated in equation 13 relates $Z[C(z_n)]$ to $C(z_n)$ and it will be used in this paper to obtain certain general relations rather than for the purpose of evaluation of particular transforms. Thus, although $C(z)$ can be evaluated by making use of equation 14, a more convenient procedure will be now illustrated by a simple example. With reference to Fig. 1(C), let $E(s)=H(s)=1/s$. It follows, from the previous discussion, that

$$E(z_n)=H(z_n)=1/1-z_n^{-1}$$

and therefore

$$C(z_n)=1/(1-z_n^{-1})^2$$

The next step is to find a function $C(s)$, such that

$$Z_n[C(s)]=C(z_n)=1/(1-z_n^{-1})^2$$

Recognizing that

$$Tz_n^{-1}/n(1-z_n^{-1})^2=Z_n[1/s^2]$$

and therefore that

$$\frac{1}{(1-z_n^{-1})^2}=\frac{nTz_n^{-1}z_n}{Tn(1-z_n^{-1})^2}=\frac{n}{T}Z_n\left[\frac{\epsilon^{sT/n}}{s^2}\right]$$

it follows that

$$Z[C(z_n)]=\frac{n}{T}Z\left[\frac{\epsilon^{sT/n}}{s^2}\right]$$

which can be determined from Appendix I. This procedure is quite general since a more complicated function $C(z_n)$ can be always expanded into partial fractions, each term of the expansion having an easily recognizable inverse.

Mathematical Description of the Multirate Controller

The program of the multirate controller is described by $D(z_n)$ which in general is equal to a ratio of two polynomials in z_n^{-1}, where as before $z_n=\epsilon^{sT/n}$, T being the sampling period of the uncompensated system. The output pulses of the multirate controller occur at the instants when $t=kT/n$ and their magnitudes are a linear combination of the past and present input and output samples. For $n=1$, $D(z_n)$ reduces to $D(z)$ which describes a program of the conventional, single-rate controller.

The whole system is shown in Fig. 2. it is seen there that the output of the controller smoothed by the zero-order hold circuit appears at the input to the plant. It is obvious that the plant is excited here by a smoother time function than in the case of a single-rate controller.

The synthesis procedure outlined in this paper follows the fundamental philosophy of the time-domain design developed by Bergen and Ragazzini.[8] The first step is to obtain an expression for $D(z_n)$ in terms of an over-all transfer function of the compensated system. From inspection of Fig. 2, three equations are obtained:

$$E_1(z) = R(z) - C(z) \tag{15}$$

$$C(z) = E_1(z)Z[D(z_n)G(s)] = E_1(z)Z[D(z_n)G(z_n)] \tag{16}$$

$$C(z_n) = E_1(z)D(z_n)G(z_n) \tag{17}$$

Eliminating $C(z)$ from equations 15 and 16 and solving for $E_1(z)$,

$$E_1(z) = \frac{R(z)}{1+Z[D(z_n)G(z_n)]} \tag{18}$$

Substituting the expression for $E_1(z)$, given by equation 18, into equation 17, the following relation is obtained:

$$C(z_n) = \frac{R(z)D(z_n)G(z_n)}{1+Z[D(z_n)G(z_n)]} \tag{19}$$

Let

$$K(z_n) \triangleq \frac{C(z_n)}{R(z)}$$

$K(z_n)$ is then the Z_n-transform of the impulsive response of the feedback system under consideration. It follows from equation 19 that $K(z_n)$ is given by

$$K(z_n) = \frac{D(z_n)G(z_n)}{1+Z[D(z_n)G(z_n)]} \tag{20}$$

Equation 20 can be rewritten as

$$Z[D(z_n)G(z_n)]K(z_n) - D(z_n)G(z_n) = -K(z_n) \tag{21}$$

Equation 21 is a functional equation of the form

$$Z[F(z_n)]K(z_n) - F(z_n) = -K(z_n) \tag{22}$$

where $F(z_n)$ is the unknown function. Solution of the foregoing equation is given by

$$F(z_n) = \frac{K(z_n)}{1-K(z)} \tag{23}$$

and since $F(z_n)$ in this case is equal to $D(z_n)G(z_n)$, the expression for $D(z_n)$ is

$$D(z_n) = \frac{1}{G(z_n)}\,\frac{K(z_n)}{1-K(z)} \tag{24}$$

where $K(z) \triangleq Z[K(z_n)]$ is the Z-transform of the impulsive response of the system shown in Fig. 2. The derivation of the solution of equation 22 is given in Appendix II. It is sufficient to observe here that equation 23 satisfies equation 22. Thus, substituting equation 23 into the left-hand side of equation 22, the resulting expression is reduced, after a few simple steps to the right-hand side of equation 22.

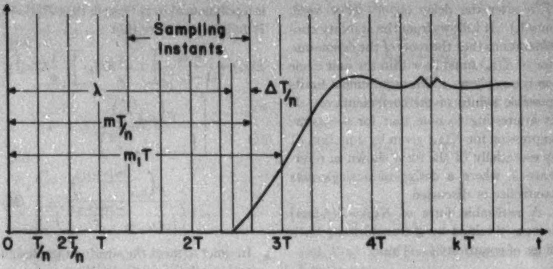


Fig. 3. Typical response of plant with lag time to step input

$$Z\left[\frac{K(z_n)}{1-K(z)}\right]K(z_n) - \frac{K(z_n)}{1-K(z)} = \frac{Z[K(z_n)]K(z_n)-K(z_n)}{1-K(z)} = -K(z_n) \tag{25}$$

It can be observed that, for $n=1$, the expression for $D(z_n)$ given by equation 24 reduces to

$$D(z) = \frac{1}{G(z)}\,\frac{K(z)}{1-K(z)} \tag{26}$$

which is identically equal to the program of a single-rate controller (reference 3).

Physical Realizability of $K(z_n)$

Equation 24 expresses $D(z_n)$ in terms of $G(z_n)$, which is assumed to be known, and the over-all pulsed transfer functions of the system, $K(z_n)$ and $K(z)$ which are to be prescribed by the designer. It is therefore of great interest to be able to specify forms of $K(z_n)$ and $K(z)$ which can be physically realized by the use of a multirate digital controller.

Physical realizability of a discrete network transfer function has been discussed by Salzer.[7] His result, applied to $D(z_n)$ and translated into the terminology of this report, states, that a physically realizable $D(z_n)$ can be written

$$D(z_n) = \frac{\sum_{k=0}^{l} a_k z_n^{-k}}{1+\sum_{k=1}^{r} b_k z_n^{-k}} \tag{27}$$

where a_k and b_k are any real numbers.

It is observed next that $G(z_n)$ can be written as

$$G(z_n) = Z_n\left[\frac{(1-\epsilon^{-sT/n})}{s}G_p(s)\right] = (1-z_n^{-1})Z_n\left[\frac{G_p(s)}{s}\right] \tag{28}$$

In many cases $G_p(s)$ can be written as $G_1(s)\epsilon^{-\lambda s}$, where $G_1(s)$ is the ratio of two polynomials in s, the degree of the denominator being higher than the degree of the numerator, and λ is an inherent delay lag. A typical response of such a plant is shown in Fig. 3, where it is seen that λ can be expressed as $\lambda = mT/n - \Delta T/n$, m being the number of the first sampling instant (sampling with the period T/n) at which the step response has a non-zero value, $0 \leq \Delta < 1$. Equation 28 can be therefore rewritten as

$$G(z_n) = z_n^{-m}(1-z_n^{-1})Z_n\left[\frac{G_1(s)}{s}\epsilon^{\Delta sT/n}\right] \tag{29}$$

or more generally as

$$G(z_n) = \frac{\sum_{k=0}^{t} g_{m+k} z_n^{-(m+k)}}{1+\sum_{k=1}^{t} h_k z_n^{-k}} \tag{30}$$

Substituting the expressions for $D(z_n)$ and $G(z_n)$, given by equations 27 and 30, in equation 20 it can be shown after some algebraic manipulations and with the help of the identity stated in equation 13 that, in general, a physically realizable $K(z_n)$ can be written as

$$K(z_n) = \frac{\sum_{\nu=0}^{l} p_{m+\nu} z_n^{-(m+\nu)}}{1+\sum_{k=1}^{s} c_k z_n^{-kn}} \tag{31}$$

where p_ν and c_k can be any real numbers and m is the number of the first interval

T/n after the delay lag λ. (For $\lambda=0$, $m=1$.) It follows from the stability considerations that the roots of the denominator of $K(z_n)$ must lie within the unit circle on the z_n-plane. This requirement limits possible values of the coefficients c_k. It is interesting to note that for $n=1$, the expression for $K(z_n)$ given by equation 31 is essentially of the form shown in reference 3, where a design of a single-rate controller is discussed.

A realizable form of $K(z)=Z(K(z_n)]$ is now obtained by Z-transforming both sides of equation 31. Thus,

$$K(z)=\frac{\sum_{\nu=0}^{q} t_{m_1+\nu} z^{-(m_1+\nu)}}{1+\sum_{k=1}^{s} c_k z^{-k}} \tag{32}$$

where m_1 is a number of the first interval T after the delay lag λ (see Fig. 2).

Restrictions Imposed by Steady-State Requirements

In addition to restrictions on the form of $K(z_n)$ and $K(z)$ which follow from the physical realizability requirement there exist constraints which are due to the required over-all performance of the control system. The design condition imposed here is that the steady-state system error be zero at the sampling instants kT/n in response to inputs whose Laplace transform has the form $R(s)=1/s^r$. The continuous system error $E(s)$ is given by $R(s)-C(s)$ and its Z_n-transform is therefore

$$E(z_n)=R(z_n)-C(z_n) \tag{33}$$

It follows from the definition of $K(z_n)$ that

$$C(z_n)=R(z)K(z_n) \tag{34}$$

Therefore $E(z_n)$ can be written as

$$E(z_n)=R(z_n)-R(z)K(z_n) \tag{35}$$

In view of equation 13, $R(z)$ can be written

$$R(z)=\frac{1}{n}\left[R(z_n)+R\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)+\ldots+R\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right] \tag{36}$$

Substituting equation 36 into equation 35 and rearranging the terms, the following relation is obtained

$$E(z_n)=R(z_n)\left[1-\frac{1}{n}K(z_n)\right]-\frac{1}{n}K(z_n)\times \left[R\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)+\ldots+R\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right] \tag{37}$$

Since the assumed form of $R(s)$ is $1/s^r$, it can be shown that $R(z_n)$ can be expressed as $P(z_n)/(1-z_n)^r$, where $P(z_n)$ is a polynomial in z_n^{-1}, such that $P(1)\neq 0$. It follows therefore that

$$E(z_n)=\frac{P(z_n)}{(1-z_n^{-1})^r}\left[1-\frac{1}{n}K(z_n)\right]-\frac{1}{n}K(z_n)\times \left[\frac{P\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)}{\left(1-z_n^{-1}\epsilon^{\frac{-j2\pi}{n}}\right)^r}+\ldots+\frac{P\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)}{\left(1-z_n^{-1}\epsilon^{\frac{-j2\pi(n-1)}{n}}\right)^r}\right] \tag{38}$$

In order to meet the steady-state design criterion all the poles of $E(z_n)$ must lie within the unit circle on the z_n-plane. It follows from equation 38 that to satisfy this condition $1-(1/n)K(z_n)$ must contain $(1-z_n^{-1})^r$ as a factor and $K(z_n)$ must also contain

$$\left[\left(1-z_n^{-1}\epsilon^{\frac{-j2\pi}{n}}\right)\left(1-z_n^{-1}\epsilon^{\frac{-j4\pi}{n}}\right)\ldots\left(1-z_n^{-1}\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right]^r$$

as a factor. This latter expression is identically equal to

$$\left[\frac{1-z_n^{-n}}{1-z_n^{-1}}\right]^r\equiv[1+z_n^{-1}+\ldots+z_n^{-(n-1)}]^r$$

In view of the the foregoing discussion and equation 31 it can be stated that $K(z_n)$ has to be such that

$$1-\frac{1}{n}K(z_n)=\frac{(1-z_n^{-1})^r A_1(z_n)}{1+\sum_{k=1}^{s} c_k z_n^{-kn}} \tag{39}$$

and

$$K(z_n)=\frac{[1+z_n^{-1}+\ldots+z_n^{-(n-1)}]^r B(z_n)}{1+\sum_{k=1}^{s} c_k z_n^{-kn}} \tag{40}$$

where $A_1(z_n)$ and $B(z_n)$ are polynomials in z_n^{-1}, whose coefficients have to be determined so that equations 39 and 40 can be satisfied simultaneously and $K(z_n)$ be of the form dictated by equation 31. The requirement that $K(z_n)$ be a stable transfer function places an obvious restriction on the permissible values of the coefficients c_k.

The assumed design criterion of the steady-state control error being zero at the instants kT/n implies that the steady-state control error must be also zero at the instants pT (k and p being integers) or as it has been already shown[3] that

$$1-K(z)=\frac{(1-z^{-1})^r A(z)}{1+\sum_{k=1}^{s} c_k z^{-k}} \tag{41}$$

where $A(z)$ is a polynomial in z^{-1}, whose coefficients have to be determined to make $K(z)$ of the form given by equation 32 and at the same time to satisfy the basic relation $Z[K(z_n)]=K(z)$. This relation interconnects equations 40 and 41. It is shown in Appendix III that these two equations contain the condition expressed by equation 39. This means that the steady-state system error is zero at the instants kT/n when either equations 39 and 40 or equations 40 and 41 are satisfied. The choice of the second set of equations simplifies algebraic manipulations and is therefore to be recommended.

Finite Settling Time Polynomials

One of the advantages gained by the use of a discrete compensator is the possibility of making the over-all response of the system settle down to a desired form within a finite time. This is done by specifying a polynomial form for the over-all transfer function.[7] To achieve the minimum settling time, these polynomials should be of the lowest degree possible. By setting coefficients c_k to zero in equations 40 and 41, the necessary design equations become

$$K(z_n)=[1+z_n^{-1}+\ldots+z_n^{-(n-1)}]^r B(z_n) \tag{42}$$

and

$$1-K(z)=(1-z^{-1})^r A(z) \tag{43}$$

If the system is to follow a step input with a zero steady-state error the exponent r is set equal to unity and for no transportation time lag ($\lambda=0$) a minimum polynomial for $K(z)$ is obtained when $A(z)=1$. It follows therefore from equation 43 that $K(z)=z^{-1}$. $B(z_n)$ can be determined now from the condition $Z[K(z_n)]=z^{-1}$. This condition implies that the minimum degree of $K(z_n)$ is n and therefore for the minimum settling time design, $B(z_n)$ must be of the form

$$B(z_n)=b_0+b_1 z_n^{-1} \tag{44}$$

Substituting equation 44 into equation 42 and remembering that $Z[K(z_n)]$ means taking every nth sample of $K(z_n)$, it follows that $b_0=0$ and $b_1=1$. Therefore in this case $K(z_n)$ must be equal to

$$K(z_n)=z_n^{-1}+z_n^{-2}+\ldots+z_n^{-n} =\frac{1-z_n^{-n}}{1-z_n^{-1}}z_n^{-1} \tag{45}$$

The fact that $b_0=0$ could have been also deduced from equation 31 by setting $m=$ for $\lambda=0$. For $\lambda\neq 0$, it can be shown that

$$K(z_n)=\frac{1-z_n^{-n}}{1-z_n^{-1}}z_n^{-m} \tag{46}$$

where m is the number of the sampling

instant (sampling period T/n) immediately following the lag time λ.

To obtain $K(z_n)$ for the case when $r>1$ it will be easier to follow a slightly modified procedure which will be now outlined in detail for the case when $r=2$, or in other words, when the system is to follow a ramp input with a steady-state error equal to zero at the sampling instants kT/n. The first step is to determine the required form of $K(z)$. Since the condition expressed by equation 43 is exactly the same as the one which is used in the design of a single-rate controller, $K(z)$ can be obtained directly from the table of prototypes given in reference 3. In general, when $r=2$ the following equation can be written

$$K(z)=(m_1+1)z^{-m_1}-m_1z^{-(m_1+1)} \tag{47}$$

where m_1 is the number of the sampling instant (sampling period T) immediately following the lag time λ. It follows from equations 31, 42, and 47 that $K(z_n)$ must be of the form

$$K(z_n)=[1+z_n^{-1}+\ldots+z_n^{-(n-1)}]^2\times [b_mz_n^{-m}+b_{m+1}z_n^{-(m+1)}]$$
$$=\frac{(1-z_n^{-n})^2}{(1-z_n^{-1})^2}[b_mz_n^{-m}+b_{m+1}z_n^{-(m+1)}] \tag{48}$$

Z-transforming both sides of equation 48, it can be written that

$$\frac{K(z)}{(1-z^{-1})^2}=\frac{b_m}{\frac{T}{n}}Z\left[\frac{\frac{T}{n}z_n^{-m}}{(1-z_n^{-1})^2}\right]+ \frac{b_m+1}{\frac{T}{n}}Z\left[\frac{\frac{T}{n}z_n^{-(m+1)}}{(1-z_n^{-1})^2}\right] \tag{49}$$

Since

$$\frac{\frac{T}{n}z_n^{-1}}{(1-z_n^{-1})^2}=Z_n\left[\frac{1}{s^2}\right] \tag{50}$$

therefore

$$Z\left[\frac{\frac{T}{n}z_n^{-m}}{(1-z_n^{-1})^2}\right]=Z\left[\frac{\epsilon^{-(m-1)sT/n}}{s^2}\right] \tag{51}$$

and

$$Z\left[\frac{\frac{T}{n}z_n^{(-m+1)}}{(1-z_n^{-1})^2}\right]=Z\left[\frac{\epsilon^{-msT/n}}{s^2}\right]$$

Observing that $-(m-1)/n=-m_1+(m_1n-m+1)/n$ and $-m/n=-m_1+(m_1n-m)/m$ where $0<(m_1n-m+1)/n<1$ and $0\leq(m_1n-m)/n<1$, it is found by means of Appendix I that

$$Z\left[\frac{\epsilon^{-(m-1)sT/n}}{s^2}\right] =\frac{\frac{m_1n-m}{n}T+\frac{m-1-n(m_1-1)}{n}Tz^{-1}}{(1-z^{-1})^2}z^{-m_1} \tag{52}$$

and

$$Z\left[\frac{\epsilon^{-msT/n}}{s^2}\right] =\frac{\frac{m_1n-m}{n}T+\frac{m-n(m_1-1)}{n}Tz^{-1}}{(1-z^{-1})^2}z^{-m_1} \tag{53}$$

In view of equations 51, 52, and 53, equation 49 can be rewritten as

$$K(z)=[b_m(m_1n-m+1)+b_{m+1}(m_1n-m)]\times z^{-m_1}+\{b_m[m-1-n(m_1-1)]+ b_{m+1}[m-n(m_1-1)]\}z^{-(m_1+1)} \tag{54}$$

Comparing the coefficients of equal powers of z^{-1} in equations 47 and 54, two simultaneous equations are obtained

$$\left.\begin{aligned} &b_m(m_1n-m+1)+b_{m+1}(m_1n-m)=m_1+1\\ &b_m[m-1-n(m_1-1)]+b_{m+1}[m-n\times (m_1-1)]=-m_1\end{aligned}\right\} \tag{55}$$

from which b_m and b_{m+1} are readily found.

A similar procedure can be adopted to determine $K(z_n)$ for $r>2$. Results of the foregoing computations are tabulated in Table I for different values of m and m_1 with the corresponding expressions for $D(z_n)G(z_n)$ obtained with equation 24.

ILLUSTRATIVE EXAMPLE

To illustrate the use of Table I let the problem be to design a controller $D(z_n)$ for a plant $G_p(s)=4.6/s+4.6$, to achieve a zero steady-state error in response to a ramp input assuming that $T=1$ second, and $n=2$. Since $\lambda=0$, therefore, $m_1=m=1$ and for a ramp input $r=2$. Corresponding expressions obtained from Table I read (for $n=2$)

$$D(z_2)G(z_2)=\frac{z_2^{-1}(1.5-z_2^{-1})}{(1-z_2^{-1})^2} \tag{56}$$

and

$$K(z_2)=(1+z_2^{-1})^2(1.5-z_2^{-1})z_2^{-1} \tag{57}$$

The whole system is shown in Fig. 4(A). It is seen there that

$$G(z_2)=Z_2\left[\frac{4.6(1-\epsilon^{-sT/2})}{s(s+4.6)}\right]$$

which becomes upon evaluation

$$G(z_2)=\frac{0.9z_2^{-1}}{(1-0.1z_2^{-1})} \tag{58}$$

Substituting equation 58 into equation 56, it is found that

$$D(z_2)=\frac{1.5-1.15z_2^{-1}+0.1z_2^{-2}}{0.9-1.8z_2^{-1}+0.9z_2^{-2}} \tag{59}$$

To compute the output $C(z_2)$ in response to a ramp input $R(s)=1/s^2$ it is observed that

$$C(z_2)=R(z)K(z_2)=R(z_2)K(z_2)$$

Since in this case,

$$R(z_2^2)=\frac{z_2^{-2}}{(1-z_2^{-2})^2}$$

and $K(z_2)$ is as given by equation 57, it follows that

$$C(z_2)=\frac{(1+z_2^{-1})^2(1.5-z_2^{-1})z_2^{-3}}{(1-z_2^{-2})^2}$$
$$=1.5z_2^{-3}+2z_2^{-4}+2.5z_2^{-5}+3z_2^{-6}+ 3.5z_2^{-7}+\ldots \tag{60}$$

The foregoing expression shows that the output samples reproduce the ramp input on and after 1.5 seconds.

For the pupose of comparison a single-rate controller will be now used to compensate the same system. The whole arrangement is shown in Fig. 4(B). The appropriate expressions for $D(z_n)G(z_n)$ and $K(z_n)$ obtained from Table I, with n set equal to unity read

$$D(z)G(z)=\frac{z^{-1}(2-z^{-1})}{(1-z^{-1})^2} \tag{61}$$

and

$$K(z)=(2-z^{-1})z^{-1} \tag{62}$$

Evaluation of $G(z)$ gives

$$G(z)=Z\left[\frac{4.6(1-\epsilon^{-sT})}{s(s+4.6)}\right]=\frac{0.99z^{-1}}{1-0.01z^{-1}} \tag{63}$$

and from equations 61 and 63 $D(z)$ is found to be

$$D(z)=\frac{2-1.01z^{-1}+0.01z^{-2}}{0.99-1.98z^{-1}+0.99z^{-2}} \tag{64}$$

The output $C(z)$ corresponding to the ramp input is

$$C(z)=R(z)K(z)=\frac{z^{-2}(2-z^{-1})}{(1-z^{-1})^2}$$
$$=2z^{-2}+3z^{-3}+4z^{-4}+\ldots \tag{65}$$

It is seen from equation 65 that on and after 2 seconds the output reproduces the ramp input at the sampling instants kT ($T=1$ second). To determine the behavior of this system between the sampling instants kT the continuous output is sampled with the sampling period $T/2$. The corresponding transform $C(z_2)$ can be written from Fig. 4(B), as

$$C(z_2)=E_2(z)Z_2\left[\frac{4.6(1-\epsilon^{-sT})}{s(s+4.6)}\right] \tag{66}$$

Table I. Minimum Settling-Time Prototypes

m₁	m	K(zₙ)	D(zₙ)G(zₙ)
1	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-1}$	$\dfrac{z_n^{-1}}{1-z_n^{-1}}$
1	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-2}$	$\dfrac{z_n^{-2}}{1-z_n^{-1}}$
1	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-3}$	$\dfrac{z_n^{-3}}{1-z_n^{-1}}$
1	2	$n+1$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-(n+1)}$	$\dfrac{z_n^{-(n+1)}}{(1-z_n^{-1})(1+z_n^{-n})}$
1	2	$n+2$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-(n+2)}$	$\dfrac{z_n^{-(n+2)}}{(1-z_n^{-1})(1+z_n^{-n})}$
1	3	$2n+1$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-(2n+1)}$	$\dfrac{z_n^{-(2n+1)}}{(1-z_n^{-1})(1+z_n^{-n}+z_n^{-2n})}$
1	3	$2n+2$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]z_n^{-(2n+2)}$	$\dfrac{z_n^{-(2n+2)}}{(1-z_n^{-1})(1+z_n^{-n}+z_n^{-2n})}$
2	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{1+n}{n}-z_n^{-1}\right)z_n^{-1}$	$\dfrac{z_n^{-1}\left(\dfrac{1+n}{n}-z_n^{-1}\right)}{(1-z_n^{-1})^2}$
2	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{2+n}{n}-\dfrac{1+n}{n}z_n^{-1}\right)z_n^{-2}$	$\dfrac{z_n^{-2}\left(\dfrac{2+n}{n}-\dfrac{1+n}{n}z_n^{-1}\right)}{(1-z_n^{-1})^2}$
2	1	$[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{3+n}{n}-\dfrac{2+n}{n}z_n^{-1}\right)z_n^{-3}$	$\dfrac{z_n^{-3}\left(\dfrac{3+n}{n}-\dfrac{2+n}{n}z_n^{-1}\right)}{(1-z_n^{-1})^2}$
2	2	$n+1$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{1+2n}{n}-2z_n^{-1}\right)z_n^{-(n+1)}$	$\dfrac{z_n^{-(n+1)}\left(\dfrac{1+2n}{n}-2z_n^{-1}\right)}{(1-z_n^{-1})^2(1+2z_n^{-n})}$
2	2	$n+2$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{2+2n}{n}-\dfrac{1+2n}{n}z_n^{-1}\right)z_n^{-(n+2)}$	$\dfrac{z_n^{-(n+2)}\left(\dfrac{2+2n}{n}-\dfrac{1+2n}{n}z_n^{-1}\right)}{(1-z_n^{-1})^2(1+2z_n^{-n})}$
2	2	$n+3$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{3+2n}{n}-\dfrac{2+2n}{n}z_n^{-1}\right)z_n^{-(n+3)}$	$\dfrac{z_n^{-(n+3)}\left(\dfrac{3+2n}{n}-\dfrac{2+2n}{n}z_n^{-1}\right)}{(1-z_n^{-1})^2(1+2z_n^{-n})}$
2	3	$2n+1$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{1+3n}{n}-3z_n^{-1}\right)z_n^{-(2n+1)}$	$\dfrac{z_n^{-(2n+1)}\left(\dfrac{1+3n}{n}-3z_n^{-1}\right)}{(1-z_n^{-1})^2(1+2z_n^{-n}+3z_n^{-2n})}$
2	3	$2n+2$ … $[1+z_n^{-1}+\dots+z_n^{-(n-1)}]^2\left(\dfrac{2+3n}{n}-\dfrac{1+3n}{n}z_n^{-1}\right)z_n^{-(2n+2)}$	$\dfrac{z_n^{-(2n+2)}\left(\dfrac{2+3n}{n}-\dfrac{1+3n}{n}z_n^{-1}\right)}{(1-z_n^{-1})^2(1+2z_n^{-n}+3z_n^{-2n})}$

Since

$$E_2(z)=E_1(z)D(z)=R(z)[1-K(z)]D(z) \quad (67)$$

it follows that

$$C(z_2)=R(z)[1-K(z)]D(z)(1-z^{-1})\times Z_2[G_p(s)/s] \quad (68)$$

The evaluation of the foregoing expression gives finally

$$C(z_2)=1.82z_2^{-3}+2z_2^{-4}+2.91z_2^{-5}+3z_2^{-6}+3.91z_2^{-7}+\dots \quad (69)$$

The outputs of both the multirate and the single-rate systems in response to a ramp input, given by equations 60 and 69 are plotted in Fig. 5. The broken line showing the continuous output from the single-rate system indicates the presence of a large amount of intersampling ripple. The line showing the continuous output from the multirate system indicates a much smoother output. Thus, Fig. 5 demonstrates plainly one advantage gained by the use of a multirate controller. By comparing equations 64 and 59 it is seen that this advantage is gained at no increase in the amount of storage elements needed.

Practical Limitations

Apart from the constraints placed on the transfer function $K(z_n)$ by the required over-all performance, there exist some practical limitations which will be now considered. One such limitation arises from the consideration of the behavior of the system when the plant parameters drift slightly from their originally determined values. This problem has been discussed already in references 3 and 4 in connection with the design of a single-rate controller. It has been found there that whenever $G(z)$ has poles or zeros on or outside the unit circle, the over-all system tends to be unstable for small drifts in their positions, unless the zeros of the specified $K(z)$ include all zeros of $G(z)$ on or outside the unit circle and the zeros of $1-K(z)$ include all poles of $G(z)$ on or outside the unit circle. Similar analysis of the multirate system shows that the actual $K(z_n)$ can become unstable when $D(z_n)$ is designed to compensate a plant such that $G(z_n)$ has poles and zeros on or outside the unit circle in the z_n-plane. Thus, let $G(z_n)$ be of the form

$$G(z_n)=\frac{\prod_p(1-a_pz_n^{-1})^lP(z_n)}{\prod_q(1-b_qz_n^{-1})^kQ(z_n)} \quad (70)$$

where $P(z_n)$ and $Q(z_n)$ are polynomials in z_n^{-1}, whose zeros lie inside the unit circle, and a_b and b_q are any numbers whose magnitudes are equal to or greater than unity. To prevent the instability of the over-all system due to small drifts in the values of a_p and b_q, $K(z_n)$ and $1-K(z)$

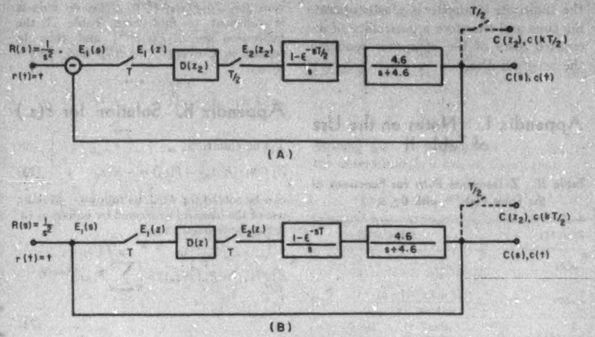


Fig. 4. Illustrative example

must be of the form of the following equations

$$K(z_n)=\prod_{p}(1-a_p z_n^{-1})^t M(z_n)\times$$

$$\prod_{q}[1+b_q z_n^{-1}+\ldots+b_q^{n-1}z_n^{-(n-1)}]^k \quad (71)$$

$$1-K(z)=\prod_{q}(1-b_q^{n}z^{-1})^k N(z) \quad (72)$$

where $M(z_n)$ and $N(z)$ are transforms which have to be determined so that equations 40 and 41 are satisfied.

Another practical limitation, which has to be considered by the designer, arises as a result of a nonlinear behavior of the plant when its input exceeds some specified value. This problem has been already discussed for the case of a single-rate system.[8] For a multirate system the situation becomes more serious with an increase in the value of the parameter n. The physical reason for this statement can be demonstrated by considering another possible advantage of multirate system, namely an improvement in the settling time of the output. For example, a minimum settling-type polynomial $K(z_n)$ obtained from Table I for a system which is to follow a step input is

$$z_n^{-1}[1+z_n^{-1}+z_n^{-2}+\ldots+z_n^{-(n-1)}]$$

the output of the system in response to the input $R(z)=1/(1-z^{-1})$ is

$$C(z_n)=z_n^{-1}+z_n^{-2}+z_n^{-3}+\ldots$$

which indicates that it takes T/n seconds to reach the step function. For an equivalent single-rate system, $K(z)=z^{-1}$ and the step function is reached in T seconds. Thus the higher the value of n, the better settling time is achieved. However, this can only be achieved at the cost of an increase in the value of the input to the plant, $G_p(s)$, during the first interval T/n.

Thus either the value of n has to be limited or the minimum settling time criterion relaxed.

General Discussion of Multirate Controller

The constraints placed on the form of $K(z)$ by equations 32 and 41 are identical to those restricting $K(z)$ in the design of a single-rate controller.[7] This means that any $K(z)$ realized by a multirate design could be also realized by the use of a single-rate controller and conversely. The multirate system behaves therefore at sampling instants kT in the same way as the equivalent single-rate system. When a single-rate system, designed on the basis of the minimum settling time, is subjected to inputs of a lower degree than those which the system was originally designed to follow, then overshoots will occur in the output before the system settles down. For example, such a system, if designed to follow a ramp input, will have a 100-per-cent overshoot in its output at $t=T$ (assuming no delay in the plant), when subjected to a step input. The physical reason for the overshoots lies mainly in the inability of the system to distinguish between, for example, a step and an advanced ramp input on the basis of the first input pulse at $t=0$. If this system was designed to follow a ramp input on the minimum settling time basis, then its output will try to reach the advanced ramp function at $t=T$, thus producing a 100-per-cent overshoot. At this moment, however, the control error produces a correcting signal and the output is reduced. Therefore, in a sampled-error system, the maximum overshoot will occur in general at some sampling instant kT, unless the plant is such as to cause a very high degree of intersampling ripple.

Similar reasoning can be applied to a multirate system. The correcting signals from the control error E_1 can be applied only at $t=kT$; at the instants the output of a multirate system is the same as the output of an equivalent single-rate system. It follows therefore that maximum overshoots in both systems have to be the same with the exception of a high ripple case when they might occur in between the sampling instants kT. Since, however, the multirate system has a tendency to minimize the intersampling ripple, therefore one can expect the overshoots in its output to be lower than in the case of a single-rate system. To sum up these arguments, the performance of a multirate system can be no worse than

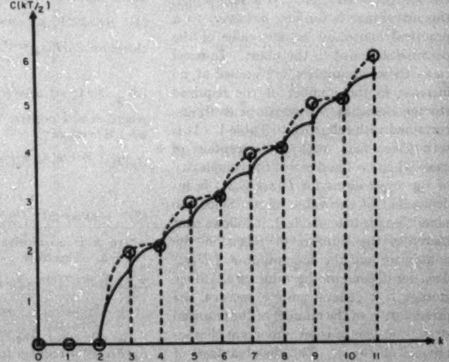


Fig. 5. Plot of outputs, c(kT/2), from systems shown in Fig. 4

- • Output of system in Fig. 4(A)
- 0 Output of system in Fig. 4(B)

that of a single-rate system as far as the overshoots are concerned.

Since the cause of the transient overshoots in both types of systems is identical, the remedy can be the same. This problem has been dealt with in some detail in reference 4. It is shown there that the overshoots can be minimized either by relaxing the condition that $K(z)$ is to be a minimum settling time polynomial or by abandoning altogether the condition of the finite settling time by assigning to coefficients c_k in equation 32 values other than zero. In both cases extra coefficients introduced into the expression for $K(z)$ are determined in such a way as to make the sum of the squared-error samples a minimum. The advantage of the second case is that when the number of the coefficients c_k is such that the degree of the denominator in the expression for $K(z)$ does not exceed the degree of the numerator, there is no increase in the number of the storage elements in the controller.

Having specified the transfer function $K(z)$ by any one of the existing methods, $K(z_n)$ can be determined from the knowledge of its general form by means of the relation $K(z)=Z[K(z_n)]$. This procedure has been already outlined in this report.

Conclusions

The multirate controller described in this paper is an example of the use of multirate methods to improve the performance of single-rate systems. The main advantage of this controller is its ability to minimize the intersampling ripple and it can be therefore used whenever there is a large ripple in the output of a system obtained by a single-rate design. An additional advantage is an improvement in the settling time when the system is subjected to a step input; this advantage is subject, however, to a practical limitation on the value of the permissible input to the plant. In most cases these advantages are gained at no increase in the number of the required storage elements. This can be easily ascertained with reference to Table I. It is seen there that, with the exception of cases when $r=1$ and $m=1$, the complexity of the expressions for $D(z_n)$ $G(z_n)$ is independent of the value of n. Since the same thing is true of $G(z_n)$, it follows that therefore the number of terms in the numerator and the denominator of $D(z_n)$ does not depend on the value of n. Since storage is in general quite expensive, the preservation of the number of the required storage elements is an important characteristic of $D(z_n)$. The disadvantage of the multirate controller is a faster operating time and therefore a possibility of increased costs of the computer realizing the program $D(z_n)$.

Appendix I. Notes on the Use of Table II

Table II. Z-Transform Pairs for Functions of the Form $F(s)\epsilon^{\Delta Ts}$ with $0\leq\Delta<1$*

$F(s)\epsilon^{\Delta Ts}$	$\Delta[F(s)\epsilon^{\Delta Ts}]$
$\frac{1}{s}\epsilon^{\Delta Ts}$	$\frac{1}{1-z^{-1}}$
$\frac{1}{s^2}\epsilon^{\Delta Ts}$	$\frac{\Delta T+Tz^{-1}(1-\Delta)}{(1-z^{-1})^2}$
$\frac{1}{s+a}\epsilon^{\Delta Ts}$	$\frac{\epsilon^{-a\Delta T}}{1-\epsilon^{-aT}z^{-1}}$
$\frac{a}{s(s+a)}\epsilon^{\Delta Ts}$	$\frac{1-\epsilon^{-a\Delta T}+z^{-1}(\epsilon^{-a\Delta T}-\epsilon^{aT})}{(1-z^{-1})(1-\epsilon^{-aT}z^{-1})}$

* Reproduced from reference 1.

To find $Z[F(s)\epsilon^{\alpha sT}$ apply the following rules:

(1) $0<\alpha<1$; set $\alpha=\Delta$

therefore $Z[F(s)\epsilon^{\alpha sT}=Z[F(s)\epsilon^{\Delta sT}]$

(2) $\alpha>1$; set $\alpha=m+\Delta$

where m is a positive integer smaller than α. Therefore

$$Z[F(s)\epsilon^{\alpha sT}]=Z[F(s)\epsilon^{msT}\epsilon^{\Delta sT}]$$
$$=z^m Z[F(s)\epsilon^{\Delta sT}]$$

(3) $-n<\alpha<0$

where n is a positive integer; set $\alpha=-n+\Delta$. Therefore

$$Z[F(s)\epsilon^{\alpha sT}]=Z[F(s)\epsilon^{-nsT}\epsilon^{\Delta sT}]$$
$$=z^{-n}Z[F(s)\epsilon^{\Delta s}]$$

Thus the problem of finding $Z[F(s)\epsilon^{\alpha sT}]$ reduces to finding the correct expression for $Z[F(s)\epsilon^{\Delta sT}]$ from Table II.

To find $Z_p[F(s)^{\alpha sT}]$ first note that it can be rewritten as $Z_p[F(s)\epsilon^{\alpha ps(T/p)}]$ then apply the following rules:

(1) $0<\alpha p<1$; set $\alpha p=\Delta$

therefore $Z_p[F(s)\epsilon^{\alpha ps(T/p)}]$
$$=Z_p[F(s)\epsilon^{\Delta s(T/p)}]$$

(2) $\alpha p>1$; set $\alpha p=m+\Delta$

where m is a positive integer smaller than αp. Therefore

$$Z_p[F(s)\epsilon^{ps(T/p)}]=Z_p[F(s)^{ms(T/p)}\epsilon^{\Delta s(T/p)}]$$
$$=z_p{}^m Z[F(s)\epsilon^{\Delta s(T/p)}]$$

(3) $-n<\alpha p<0$

where n is a positive integer; set $\alpha p=-n+\Delta$. Therefore

$$Z_p[F(s)\epsilon^{\alpha ps(T/p)}]=Z_p[F(s)\epsilon^{-ns(T/p)}\epsilon^{\Delta s(T/p)}$$
$$=z_p{}^{-n}Z_p[F(s)\epsilon^{\Delta s(T/p)}]$$

In all cases the problem of finding $Z_p[F(s)\epsilon^{\alpha sT}]$ reduces to finding the expression for $Zp[F(s)\epsilon^{\Delta(sT/p)}]$. To do this it is sufficient to find from Table II the expression for $Z[F(s)\epsilon^{\Delta sT}]$ and then to replace in it every T by T/p and every z by z_p.

Appendix II. Solution for F(z$_n$)

The equation

$$Z[F(z_n)]K(z_n)-F(z_n)=-K(z_n) \qquad (73)$$

can be solved for $F(z_n)$ as follows. Making use of the identity expressed by equation 13, it can be written that

$$Z[F(z_n)]=P_n[F(z_n)]=\frac{1}{n}\sum_{k=1}^{n}F(z_n\epsilon^{j2\pi k/n}) \qquad (74)$$

and therefore equation 73 can be rewritten as

$$\left[\frac{1}{n}\sum_{k=1}^{n}F(z_n\epsilon^{j2\pi k/n})\right]K(z_n)-F(z_n)$$
$$=-K(z_n) \qquad (75)$$

Expanding the summation and rearranging the terms in equation 75, the following is obtained:

$$F(z_n)\left[\frac{1}{n}K(z_n)-1\right]+F(z_n\epsilon^{j2\pi/n})\times$$
$$\left[\frac{1}{n}K(z_n)\right]+\ldots+F(z_n\epsilon^{j2\pi(n-1)/n})\times$$
$$\left[\frac{1}{n}K(z_n)\right]=-K(z_n) \qquad (76)$$

If in place of z_n in equation 76, $z_n\epsilon^{j2\pi/n}$, $z_n\epsilon^{j4\pi/n}$, $\ldots$, $z_n\epsilon^{j2\pi(n-1)/n}$ are substituted one after another, then a set of n algebraic equations

$$F(z_n)\left[\frac{1}{n}K(z_n)-1\right]+F\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\left[\frac{1}{n}K(z_n)\right]+$$
$$\ldots+F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K(z_n)\right] \qquad (77)$$

$$F\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)-1\right]+\ldots+$$
$$F\left(z_n\epsilon^{\frac{j2\pi}{n}}\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{-j2\pi}{n}}\right)\right]$$
$$=-K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)-1\right]+$$
$$\ldots+F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\times$$
$$\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right]=-K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)$$

with n unknowns $F(z_n)$, $F(z_n\epsilon^{j2\pi/n})$, $\ldots$, $F(z_n\epsilon^{j2\pi(n-1)/n})$ is obtained. $F(z_n)$ is the only unknown function which is of interest here. To find an expression for $F(z_n)$, the set of equations 77 is rewritten as

$$F(z_n)\left[\frac{1}{n}K(z_n)-1\right]+F\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\left[\frac{1}{n}K(z_n)\right]+\ldots+F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K(z_n)\right]=K(z_n)$$

$$F(z_n)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)-1\right]+F\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\times\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\right]+\ldots+F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\right]=-K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right) \qquad (78)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$F(z_n)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)-1\right]+F\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\times\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right]+\ldots+F\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\left[\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right]=-K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)$$

and the Cramer rule is directly applied. Thus, $F(z_n)$ can be written as

$$F(z_n)=\frac{|N|}{|D|} \qquad (79)$$

where $|N|$ and $|D|$ are determinants shown in the following:

$$|N|=\left\|\begin{matrix}-K(z_n) & \frac{1}{n}K(z_n) & \ldots\ldots & \frac{1}{n}K(z_n)\\ -K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right) & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)-1 & \ldots\ldots & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ -K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right) & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right) & \ldots\ldots & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)-1\end{matrix}\right\|$$

$$|D|=\left\|\begin{matrix}\frac{1}{n}K(z_n)-1 & \frac{1}{n}K(z_n) & \ldots\ldots & \frac{1}{n}K(z_n)\\ \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right) & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi n}{n}}\right)-1 & \ldots\ldots & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi n}{n}}\right)\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right) & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right) & \ldots\ldots & \frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)-1\end{matrix}\right\|$$

To simplify the foregoing expressions, let

$$A_1=\frac{1}{n}K(z_n), A_2=\frac{1}{n}K(z_n\epsilon^{j2\pi/n}),\ \ldots \text{ and } A_n=\frac{1}{n}K(z_n\epsilon^{j2\pi(n-1)/n}),$$

$|N|$ and $|D|$ can be then written as

$$|N|=n\left\|\begin{matrix}-A_1 & A_1 & A_1\ldots\ldots A_1\\ -A_2 & A_2-1 & A_2\ldots\ldots A_2\\ -A_3 & A_3 & A_3-1\ldots A_3\\ \cdot & \cdot & \cdot\quad\cdot\\ \cdot & \cdot & \cdot\quad\cdot\\ -A_n & A_n & A_n\ldots\ldots A_n-1\end{matrix}\right\| \text{ and } |D|=\left\|\begin{matrix}A_1-1 & A_1 & A_1\ldots\ldots A_1\\ A_2 & A_2-1 & A_2\ldots\ldots A_2\\ A_3 & A_3 & A_3-1\ldots A_3\\ \cdot & \cdot & \cdot\quad\cdot\\ \cdot & \cdot & \cdot\quad\cdot\\ A_n & A_n & A_n\ldots\ldots A_n-1\end{matrix}\right\| \qquad (80)$$

Making use of the standard properties of the determinants, it can be shown that

$$|N|=n\begin{vmatrix}-A_1 & 0 & 0\ldots & 0\\ 0 & -1 & 0\ldots & 0\\ 0 & 0 & -1\ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0\ldots & -1\end{vmatrix}=nA_1(-1)^n=\left[n\frac{1}{n}K(z_n)\right](-1)^n$$

$$=K(z_n)(-1)^n \qquad (81)$$

and

$$|D|=\left\|\begin{matrix}-1 & 0 & 0\ldots & 0\\ 0 & -1 & 0\ldots & 0\\ 0 & 0 & -1\ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0\ldots & +1\end{matrix}\right\|\left[\sum_{k=1}^{n}A_k-1\right]=(-1)^{n-1}\left[\sum_{k=1}^{n}A_k-1\right]$$

$$=(-1)^{n-1}\left[\sum_{k=1}^{n}\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)-1\right]$$

$$=\{|1-Z[K(z_n)]|\}(-1)^n$$

$$=[1-K(z)](-1)^n \qquad (82)$$

Substituting results from equations 81 and 82 into equation 79 and simplifying the expression thus obtained, the final form of $F(z_n)$ becomes

$$F(z_n)=\frac{K(z_n)}{1-K(z)} \qquad (83)$$

Appendix III. Proof of Sufficiency of Equation 39

The purpose of this appendix is to show that equations

$$K(z_n)=\frac{[1+z_n^{-1}+z_n^{-2}+\ldots+z_n^{-1(n-1)r}B(z_n)]}{1+\sum_{k=1}^{s}c_kz_n^{-kn}} \qquad (40)$$

and

$$1-K(z)=\frac{(1-z^{-1})^rA(z)}{1+\sum_{k=1}^{s}c_kz^{-k}} \qquad (41)$$

imply that

$$1-\frac{1}{n}K(z_n)=\frac{(1-z_n^{-1})^rA_1(z_n)}{1+\sum_{k=1}^{s}c_kz_n^{-kn}} \qquad (39)$$

Since $K(z)$ can be written as

$$K(z)=\frac{1}{n}K(z_n)+\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)+\ldots+\frac{1}{n}K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right) \qquad (84)$$

therefore equation 41 can be rewritten as

$$1-\frac{1}{n}K(z_n)=\frac{(1-z^{-1})A(z)}{1+\sum_{k=1}^{s}c_kz^{-k}}+\frac{1}{n}\left[K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)+\ldots+K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)\right] \qquad (85)$$

Since

$$[1+z_n^{-1}+z_n^{-2}+\ldots+z_n^{-(n-1)}]^r = \frac{(1-z_n^{-n})^r}{(1-z_n^{-1})^r} \quad (86)$$

it follows from equation 40 that

$$K\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)=\frac{(1-z_n^{-n})^r B\left(z_n\epsilon^{\frac{j2\pi}{n}}\right)}{\left(1+\sum_{k=1}^{s}c_k z_n^{-kn}\right)\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi}{n}}\right)^r}$$

$$\vdots \qquad (87)$$

$$K\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)=\frac{(1-z_n^{-n})^r B\left(z_n\epsilon^{\frac{j2\pi(n-1)}{n}}\right)}{\left(1+\sum_{k=1}^{s}c_k z_n^{-kn}\right)\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi(n-1)}{n}}\right)^r}$$

Substituting equation 87 into equation 85 and bringing all terms into a common denominator, there results the equation at the top of this page:

$$1-\frac{1}{n}K(z_n)=\frac{(1-z^{-1})^r C(z_n)}{\left[1+\sum_{k=1}^{s}c_k z_n^{-kn}\right]\left[\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi}{n}}\right)\ldots\ldots\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi(n-1)}{n}}\right)\right]} \quad (88)$$

where $C(z_n)$ is a polynomial in z_n^{-1}. To complete the proof it is sufficient to note that (see below, bottom of column)

$$\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi}{n}}\right)\ldots\left(1-z_n^{-1}\epsilon^{-\frac{j2\pi(n-1)}{n}}\right) = \frac{(1-z^{-1})^r}{(1-z_n^{-1})^r}$$

---

and therefore $C(z_n)$ must have real coefficients since $K(z_n)$ is a real function of z_n^{-1} and is specified as such by equation 40.

References

1. The Theory of Pulse Monitored Servomechanisms and Their Use for Prediction, R. H. Barker. *Report No. 1046*, Signal Research and Development Establishment, Hants, England, Nov. 1950.

2. The Pulse Transfer Function and Its Application to Sampling Servo Systems, R. H. Barker. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. IV, Dec. 1952, pp. 302–17.

3. Sampled-Data Processing Techniques for Feedback Control Systems, Arthur R. Bergen, John R. Ragazzini. *AIEE Transactions*, vol. 73, pt. II, Nov. 1954, pp. 236–47.

4. Factors in Design of Digital Controllers for Sampled-Data Feedback Systems, J. E. Bertram. *Technical Report T-8/B*, Electronics Research Laboratories, Columbia University Engineering Center, New York, N. Y., Aug. 15, 1955.

5. The Analysis of Multiple-Rate-Sampled Systems, G. M. Kranc. *Technical Report T-11/B. Ibid.*, Nov. 15, 1955.

6. Network Compensation of Error-Sampled Feedback Systems, J. Sklansky. *Doctoral Dissertation*, Department of Electrical Engineering, Columbia University, 1955.

7. Frequency Analysis of Digital Computers in Real Time, J. M. Salzer. *Proceedings*, Institute of Radio Engineers, New York, N. Y., Feb. 1954

◆

Discussion

J. E. Bertram (Columbia University, New York, N. Y.): This paper presents a fundamentally new concept for the time-domain synthesis of error-sampled feedback control systems. But as with previous methods it insures desirable performance only at sampling instants and little or no control is exercised over the value of the continuous output between sampling instants, other than that obtained by the choice of sampling rate. Certainly, in many cases this intersample ripple may be highly undesirable.

Barker[1] first suggested a method for the analysis of this intersample ripple and developed a table of modified Z-transforms to facilitate the analysis. Jury[2] has further exploited this work and has added some minor design considerations. Ragazzini and Slansky[3] have suggested another method of analysis by which the complete intersample response between the nth and $(n+1)$th samples may be evaluated. While all of these methods have useful applications, none of them are directly concerned with the synthesis of low-ripple systems.

The author in this paper has shown that the intersample ripple is reduced if the controller or discrete compensating filter is operated at a rate which is some integral multiple of the basic error-sampling rate. In this way the quantization in time is made finer. However, it should be understood that while the ripple is usually greatly reduced by this procedure, the actual control of the output is still only at discrete points in time and not over all time.

This discussion extends a technique proposed by Kalman[4] to dead-beat error-sampled feedback control systems after a finite length of time for a step input. The problem solved by Kalman can be best explained by refering to Fig. 2 and setting $n=1$. With this system it is desired to determine a $D^*(Z)$ such that upon application of a step input the output goes to this same level in the shortest possible time and remains there, not at just the sampling instants, but at all instants, until the input is again changed. That is, following a step input the system error, $e_1(t)$, should go to zero in a finite length of time and remain there until the system is again disturbed.

An extension of this technique in combination with the multirate controller has the advantage of "dead-beating" in response to a step input in one sample period at the basic error rate and requires an extremely simple controller.

For an explanation of the design procedure the system in Fig. 2 may be considered. It is possible to make this dead beat to a step input in T seconds where T is the basic error-sampling rate.

Design Algorithm

Set n equal to the highest power of Z_n^{-1} in the denominator of the Z_n-transform of the plant-hold combination. This pulse transfer function is of the form

$$Z_n\left[\frac{(1-e^{-sT/n})}{s}G_p(s)\right]=\frac{\sum_{i=1}^{n}\alpha_i Z_d^{-i}}{1+\sum_{i=1}^{n}\beta_i Z_n^{-i}}$$

The form of the multirate controller is then given by the following relation

$$D(Z_n)=\frac{1+\sum_{i=1}^{n}\beta_i Z_n^{-1}}{(1-Z_n^{-1})\sum_{i=1}^{n}\alpha_i}$$

where common factors in the numerator and denominator must be cancelled.

To show that this design does the job, assume that a step input of magnitude U is applied to the relaxed system at time $t=0$. Thus, $e_1(0)=U$ is applied to the multirate controller. Since the Z_n pulse transfer function of the controller and plant-hold combination is now

$$D(Z_n)G(Z_n)=\frac{\sum_{i=1}^{n}\alpha_i Z_n^{-i}}{\left(\sum_{i=1}^{n}\alpha_i\right)(1-Z_n^{-1})}$$

it is apparent that the output

$$c\left(\frac{nT}{n}\right)=\frac{U}{\left(\sum_{i=1}^{n}\alpha_i\right)}[\alpha_1+\alpha_2+\ldots+\alpha_n]=U$$

Therefore, at time $t=T$ when the basic-error switch closes for the second time, $e_1(T)=0$. This, of course, does not insure that the output is U for time other than sampling instants. However, if in addition the input to the plant-hold combination can be shown to be constant while the output is constant, it is then impossible for the output to change between sampling instants except in the special case where the plant-hold combination has an impulse response which is purely oscillatory and of the same frequency or some multiple of $n2\pi/T$. An inspection of the relation for $D(Z_n)$ shows that this condition is satisfied. Note that if the plant has an integration then $D(Z_n)$ is a finite polynomial in Z_n^{-1} and the plant has no further inputs after $t=T$, while if the plant has no integration $D(Z_n)$ has an integration and the input after $t=T$ remains constant at some value other than zero.

To illustrate the use of the algorithm consider an example problem. Let $T=1$ second and the plant-hold combination be given by $1-e^{-sT/n}/s\cdot 1/s(s+1)$ where the hold is changed to conform to the multirate system. By the algorithm we set $n=2$ so that

$$Z_2\left[\frac{1-e^{-sT/2}}{s}\cdot\frac{1}{s(s+1)}\right]$$
$$=\frac{e^{-1/2}Z_n^{-1}+(1-2e^{-1/2})Z_2^{-2}}{(1-Z_2^{-1})(1-e^{-1/2}Z_2^{-1})}$$

Therefore, the multirate controller is given by

$$D(Z_n)=\frac{1-e^{-1/2}Z_2^{-1}}{(1-e^{-1/2})}$$

This system responds with dead-beat performance to a step input in one second and the controller requires only one storage element. Thus with the multirate controller it is possible to obtain ripplefree performance with a settling time equal to the basic error-rate sample period.

REFERENCES

1. See reference 2 of the paper.

2. SYNTHESIS AND CRITICAL STUDY OF SAMPLED-DATA CONTROL SYSTEMS, Eliahu I. Jury. *AIEE Transactions*, vol. 75, pt. II, July 1956, pp. 141–51.

3. ANALYSIS OF ERRORS IN SAMPLED-DATA FEEDBACK SYSTEMS, Jack Sklansky, J. R. Ragazzini. *Ibid.*, vol. 74, pt. II, May 1955, pp. 65–71.

4. *Discussion* by R. E. Kalman of reference 3 of the paper, p. 245.

George M. Kranc: I wish to thank Mr. Bertram for his highly constructive contribution. The extension of Kalman's "dead-beat" method of obtaining ripplefree responses to the case of the multirate controller yields extremely promising results. This method should be further investigated to get some information about the performance of the system when the inputs are arbitrary. The design of the multirate controller for the ripplefree response to the higher order polynomial inputs seems to present no formal difficulties. The same statement applies to plants with inherent delay lags.

It seems that the main limitation of the design procedure suggested by Mr. Bertram is that it will lead to an unstable over-all system function $K(z_n)$, whenever the poles of the plant-hold combination lie on or on the outside of the unit circle on the z_n-plane (with an exception of the pole at $z_n=1$, providing its order is no greater than the order of the polynomial input for which the system is designed). In addition it is necessary to note that although $K(z_n)$ is of the desirable polynomial form, it can no longer be of the minimum-settling type. Thus the rise time is sacrificed for the ripplefree performance.

Another possible disadvantage is that the value of the parameter n may be unduly high for higher order plants. Thus the advantage of having less storage in the multirate controller may be offset by the requirement of the faster operation.

A Transient Stability and Voltage Study for a Modern Oil Refinery Distribution System

C. W. BOICE
ASSOCIATE MEMBER AIEE

S. R. DURAND
MEMBER AIEE

D. DALASTA
MEMBER AIEE

MODERN industrial power systems are being provided with electrical distribution just as advanced in planning and design as the processes which are served. Expenditures for switching and protective facilities and for automatic operating features ensure against prolonged outages and result in over-all long-range economies. Prolonged outages in most new industrial power systems are no longer common occurrences; in fact they can be considered a rarity. Yet, there often remains a question as to the adequacy of over-all continuity of the electric system of modern continuous-process plants or oil refineries.

In continuous-process plants there is an interrelation between several processes in a chain of operations. The electric power distribution system to each individual process may appear to be designed for a high degree of reliability. However, an analysis of transient stability and voltage conditions throughout the entire system during and immediately after a severe fault condition may reveal inadequacies in design for automatic recovery of power to one or more critical process units.

This paper will describe a study to determine the performance of a specific distribution system for an oil refinery. The system studied was that of the new Delaware Flying "A" Refinery which was constructed for Tide Water Oil Company. The results of an initial study served to confirm that these changes improved the system performance under faulty conditions without adding excessively to the over-all cost.

Method of Analysis

In transient-stability studies[1] involving induction machines, network calculator or digital computer methods are generally employed. The a-c network calculator was used for this study. For suitable accuracy, when induction machines are involved, the time interval in the step-by-step calculations must be kept small. Calculations of the change of angle of each motor's internal voltage for each step are lengthy, and hence time consuming. Combination network analyzer and digital computer methods can be used, the network solution for initial conditions and admittance constants being obtained with the network calculator while the remainder of the work is done digitally. Since system admittances change when machines are removed from the system (because of contactor dropout), digital computer methods have certain disadvantages for industrial power system studies. The network calculator method allows greater flexibility during the course of each run and also gives a better physical picture to those present.

Stability Problem in Oil Refineries

A shutdown and start-up of an oil refinery unit due to an electrical failure can be a very expensive operation. In general, individual units in an oil refinery are not independent since one may depend on other units for feed material

---

Paper 56-767, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer and Pacific General Meeting, San Francisco, Calif., June 25–29, 1956. Manuscript submitted March 29, 1956; made available for printing January 8, 1957.

C. W. BOICE is with C. F. Braun and Company, Alhambra, Calif., and S. R. DURAND and D. DALASTA are with Allis-Chalmers Manufacturing Company, San Francisco, Calif., and Milwaukee, Wis., respectively.

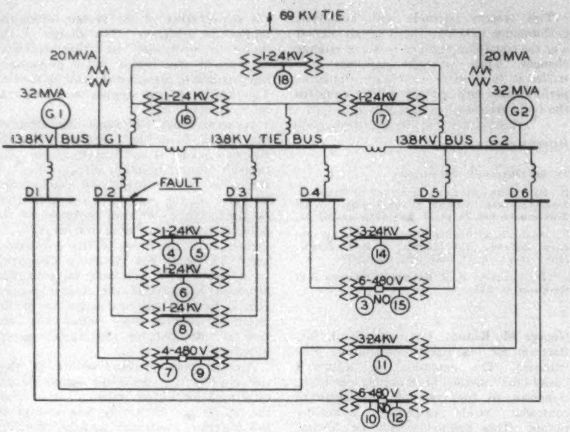


Fig. 1. Distribution system under study. Numbered circles indicate equivalent induction machines

and in turn feeds a third unit. Depending on the length of time required to restart, a forced shutdown of a single unit can affect the production of the entire refinery. Therefore a power failure of more than a few seconds' duration to one process in the chain may necessitate a costly shutdown or curtailment of production of all other processes.

The transient stability of a modern oil refinery power system is a matter of prime importance. This is particularly true when a large proportion of the motor starters throughout the refinery are equipped with time delay undervoltage protection. Time delay adjustments are usually set so that after a system disturbance, motor starters that have dropped out because of low voltage will automatically reclose if adequate voltage recovery takes place within 3 to 4 seconds. With this type of control, the following three questions with relation to the transient stability of the system must be answered:

1. Can the power system selected handle the simultaneous reacceleration currents of large numbers of motors connected to each substation?
2. Can all motors reaccelerate their loads under conditions existing immediately after the clearing of a severe fault on the system?
3. If not, to what extent throughout the refinery would motor loads be adversely affected by faults at critical points in the system?

DESCRIPTION OF SYSTEM

The refinery distribution system is shown in Fig. 1. Each of two 13.8-kv generator busses is supplied by one 32-megavolt-ampere generator, a 20-megavolt-ampere utility tie transformer, and a reactor tie to the 13.8-kv tie bus. These busses feed six 13.8-kv distribution busses through reactors. Each transformer of the 2,400-volt and 480-volt double-ended unit substations is supplied by 13.8-kv underground distribution cables from feeder circuit breakers connected to different generator busses. The twelve 2,400-volt substations are rated 10,000 to 12,500 kva each and are of the spot network type. The sixteen 480-volt substations rated 1,500 kva each are designed for low-voltage selective operation with normally open-tie breakers. Motors *16*, *17*, and *18* represent station auxiliaries and are supplied through reactors off the main busses at the generating station.

Most of the refinery motors are squirrel-cage induction-type driving centrifugal pumps. A few large slow-speed induction motors drive reciprocating compressors. Each 2,400-volt substation load was represented by an equivalent machine. Because of the normally open-tie circuit, each 480-volt substation required two equivalent machines. With one 2,400-volt substation supplying several reciprocating drives as well as some centrifugal drives, it was desirable to use two equivalent machines at that particular location. The total number of equivalent machines set up in the a-c network board was fifteen induction motors, two steam turbine generators, and the power company tie line.

TYPE OF DISTURBANCE UNDER STUDY

A choice had to be made relative to the point of origin of the disturbance and its nature. This selection was important

Table I. Induction Motors Represented by Equivalent Machines

Equivalent Motor	Voltage	Type of Drive	Torque Characteristic	No. of Motors and Horsepower	RPM
4	2,300	rcp*	high†	1 .. 1,000 1 .. 800 2 .. 400 1 .. 350 1 .. 200	400 400 514 514 514
5	2,300	cent‡	normal	1 .. 800 3 .. 500 1 .. 150 1 .. 200	3,600 3,600 3,600 1,800
6	2,300	rcp	high*	2 .. 3,000	327
7, 9	440	cent	normal	9 .. 125 8 .. 60 16 .. 7½ 72 .. 2 4 .. 60 10 .. 7½	1,800 1,800 1,800 1,800 3,600 3,600
8	2,300	cent	normal	3 .. 2,000	600
10, 12, 13, 15	440	same as motor 7, except 150% quantity			
11, 14	2,300	cent	normal	5 .. 800 2 .. 500 4 .. 250 2 .. 150 4 .. 300 9 .. 200 3 .. 2,000 1 .. 150	3,600 3,600 3,600 3,600 1,800 1,800 600 400
16, 17, 18	2,300	cent	normal	2 .. 900 3 .. 500 1 .. 1,250 1 .. 450	3,600 1,800 1,800 450

* Reciprocating. † Normal for initial study. ‡ Centrifugal.

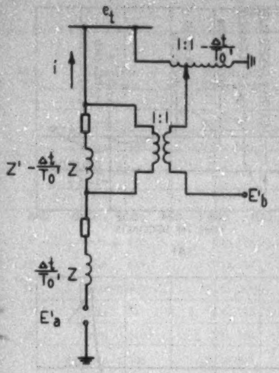

e_t
i
$1:1 - \frac{\Delta t}{T_0'}$
1:1
$Z' - \frac{\Delta t}{T_0'} Z$
E'_b
$\frac{\Delta t}{T_0'} Z$
E'_a

Fig. 2. Induction motor transient equivalent circuit including analog computing circuit

$$e_b' = e_t\left(1 - \frac{\Delta t}{T_0'}\right) + 1\left(Z' - \frac{\Delta t}{T_0'} Z\right)$$

$$Z = r_s + jx$$
$$Z' = r_s + jx'$$
$$X = X_s + X_m$$
$$X' = X_s + \frac{X_r X_m}{X_r + X_m}$$

X_s = leakage reactance of stator
X_r = leakage reactance of rotor
X_m = magnetizing branch reactance
r_s = induction machine a-c stator resistance
Δt = time interval, seconds
i = stator current
T_0' = open circuit time constant, seconds
E_a' = internal voltage at beginning of interval
E_b' = internal voltage at end of interval
e_t = terminal voltage

since, if properly made, only one or two conditions would need to be investigated to determine satisfactorily the system performance. An electrical failure either at one of the 13.8-kv generator terminals or at one of the 13.8-kv distribution busses seemed most logical. Because of segregated phase construction of switchgear, bus ducts, etc., at the generator busses, faults at that point involving more than a single phase to ground would be remote. The switchgear and connections at the 13.8-kv distribution busses, however were planned for more economical nonsegregated construction, and phase-to-phase faults can develop. This consideration led to the assumption that the most severe system disturbance to be studied was a solid 3-phase fault at a 13.8-kv distribution bus, *D2*, nearest the substations containing the reciprocating drives. It was felt that this choice would best depict the over-all system performance because of the very symmetry of the arrangement.

EQUIVALENT MACHINE REPRESENTATION

At some substations the only load consists of several large motors of identical design and size for which an equivalent motor was easily determined. At other substations for both 2,400 and 480 volts the motors vary greatly both in size and speed. In some cases the method of weighted averages was used to determine the equivalent machine. For the 480-volt substations it became more practical to choose three typical machines and combine them in the same manner. The philosophy behind the methods of obtaining equivalents was as follows: if an equivalent was required to properly portray the motors' effect on the system as a whole, the method of weighted averages was used. On the other hand, if the performance of the motor itself is of considerable interest, then it became necessary to allow the equivalent to approach the characteristics of a particular motor. Some compromise was necessary. However, in general, the former method was used for the machines electrically remote from the fault and the latter for those adjacent to the fault. The list of induction motors which are being represented by equivalent machines is given in Table I.

PROCEDURE

With the positive sequence impedances set up on the network calculator and the initial conditions known, the motor transient stability equivalents, Fig. 2, were set in and each motor's internal voltage was adjusted to give the proper kilowatt and kilovar flow. Loads were arbitrarily increased about 10%, partially to compensate for the error involved in the use of 0.02-second intervals and to ensure conservative results. Step-by-step calculations of motor internal angle changes were made after each set of readings. The new angles and the new internal voltage E_b' for each induction motor were set in for the next interval at point E_a' of Fig. 2. The generator internal voltages were held constant for the period of the study and the generator internal angle changes calculated in the normal manner.[3] This procedure was carried out until the motor slips and power inputs pretty well indicated whether the motor would reaccelerate to initial speed or would stall. The fault-clearing time in the study in question was assumed to be 0.11 second or roughly $6^1/_2$ cycles. The 13.8-kv circuit-breaker interrupting time would be about 3 or 4 cycles and the remaining period was allowed for relay time.

Table II. Assumed Performance Characteristics of Motor Control Contactors

Stem	Type of Control 480 Volts, Standard	 2,400 Volts, Standard	 2,400 Volts, Time Delay
Voltages in per-cent rated			
average pickup	75	75	75
average dropout, first study	40	40	none
average dropout, second study	50	30	30
Time from dropout voltage, cycles			
until contacts part	*	1	20
until arc is interrupted	*	2	21
Time from pickup voltage, cycles, until contacts touch	3	13	13

Circuit breaker control mechanically latched in; adjustable time delay undervoltage tripping.

* Time involved negligible.

VOLTAGE STUDY

Inasmuch as short-circuit studies to determine momentary and interrupting capacities of circuit breakers and settings of overcurrent relays had already been made, fault currents were of no special concern. However, the transient voltages at all main and substation busses were of particular interest. During a disturbance, the voltages around the system drop, in many cases sufficiently low to allow contactor dropout. Circuit-breaker control had been planned for station auxiliaries as well as most other motors of over 1,000 horsepower. For the bulk of the motors, standard motor control units were planned. The 2,400-volt units were designed for standard air contactors with fuses of high interrupting capacity, while the 480-volt starters consisted of combination units with molded case circuit breakers installed in control centers.

Table II shows the average performance characteristics of the various control devices:

1. 480-volt contactors.
2. 2,400-volt contactors, standard dropout.
3. 2,400-volt contactors, time delay dropout.
4. Circuit-breaker control.

These data were used to drop off and pick up various motors as the motor bus voltage fell upon occurrence of the fault, and recovered when the fault was cleared. Whenever a motor was disconnected from the system due to voltage conditions, it was found convenient to carry out the normal calculations of motor internal angle and to read the internal voltage magnitude for succeeding intervals. This took little additional effort

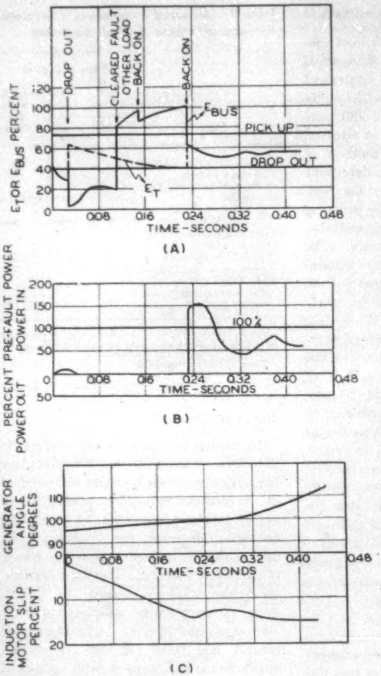


Fig. 3 (left). Performance curves of motor 4, initial study

A—Bus and motor terminal voltage
B—Prefault power
C—Motor slip and generator angle

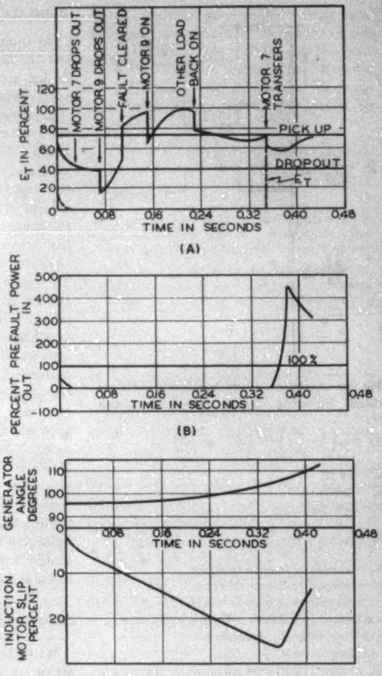


Fig. 4 (right). Performance curves of motor 7, initial study

A—Terminal voltage for motor 7 (dashed line) and bus voltage for motor 9 (solid line)
B—Per-cent prefault power
C—Motor slip and generator angle

but kept the machine in readiness to be restored to the system as soon as the voltage recovered to the pickup value and the contactor had time to close. During the disconnected period the motor internal voltage followed its open-circuit decay curve. Reconnecting the 2,400-volt machines to the system nearly always caused a rather severe voltage reduction. This was due to the relatively long dropout and pickup times which kept these motors disconnected from the system for a definite minimum time even if the motor bus voltage was restored at the very instant of dropout.

Results of Initial Study

Performance curves for one equivalent 2,400-volt machine (motor *4*) are shown in Fig. 3. At the very instant of fault (time 0+), the voltage at the motor bus dropped to about 40% causing the contactor to start to out immediately. At about 0.03 second, the contactor isolates the motor and the motor voltage (residual, dashed line) jumps up while the motor bus voltage is nearly zero (fault near-by). The fault is cleared at bus *D2*, Fig. 1, in 0.11 second by differential relay operation causing all breakers to be tripped, clearing the faulted bus. Although the motor bus voltage is restored to 85% immediately from bus *D3*, the motor is not reconnected by the closing of its contactor until 0.23 second, at which time its open-circuit voltage has decayed to about 40% from nearly 70% at 0.03 second. The motor reaccelerating current again pulls the bus voltage down upon reconnection, this time to about 60%, but not far enough to cause a second drop out of the motor contactor. However, it does pull the voltage down so far that the per-cent prefault power into the motor remains well under 100%, thus preventing reacceleration.

The per-cent of slip curve shows slowly increasing slip after a feeble attempt at recovery from 0.24 to 0.32 second. The generator internal angles had advanced to about 20 degrees during the course of the run. Since the motor is likely to stall, some means of preventing or delaying the contactor dropout or providing higher motor torque characteristics should be considered.

This is an example of the worst effect of a bus fault on a motor supplied from a substation that is directly connected to the faulted bus.

The performance curves for a typical 440-volt substation are shown in Fig. 4. The bus voltage is for motor *9* (connected to unfaulted side of normally open-tie breaker) while the prefault power and slip curves are for motor *7* (connected to side nearest fault). The sequence and magnitude of abrupt voltage variations are clearly evident. Motor *7* had reached a maximum slip of about 25% at the time the automatic transfer was completed but recovered when reconnected to the line, the voltage dropping to slightly below 60% but well short of the 40% dropout value.

Table III. Results of Confirming Study for Conditions A and B

Motor No.	Type of Drive Voltage	Size Range, Horsepower	Starting Torque Std.* or High	Control	Study	Minimum Value Bus Voltage, Per Cent	Discharge from Bus	Time-off Line, Seconds	Maximum Value Slip, Per Cent
4	rcp† 2,300	200 to 1,000	high	19-cycle drop‡	A	18	no		8
					B	15	no		8
5	cent§ 2,300	150 to 800	std	std drop	A	18	yes	0.22	8
					B	15	yes	0.30	10
6	rcp 2,300	3,000	high	circuit breaker	A	18	no		6
					B	15	no		7
7	cent 440	2 to 125	std	std drop	A	0	yes		100
					B	0	yes		100
8	cent 2,300	2,000	std	circuit breaker	A	18	no		6
					B	14	no		7
9	cent 440	2 to 125	std	std drop	A	37	yes	0.14	13
					B	30	yes	0.28	18
10	cent 440	2 to 125	std	std drop	A	37	yes	0.12	11
					B	33	yes	0.20	14
11	cent 2,300	150 to 2,000	std	19-cycle drop	A	43	no		3
					B	34	no		4
12	cent 440	2 to 125	std	std drop	A	50	no		6
					B	38	yes	0.20	14
13	cent 440	2 to 125	std	std drop	A	44	yes	0.14	14
					B	34	yes	0.22	15
14	cent 2,300	150 to 800	std	19-cycle drop	A	46	no		3
					B	35	no		4
15	cent 440	2 to 125	std	std drop	A	50	no		6
					B	38	yes	0.22	15
16	cent 2,300	450 to 1,250	std	circuit breaker	A	40	no		3
					B	32	no		4
17	cent 2,300	450 to 1,250	std	circuit breaker	A	47	no		3
					B	35	no		4
18	cent 2,300	450 to 1,250	std	circuit breaker	A	44	no		3
					B	35	no		3

* Standard. † Reciprocating. ‡ Dropout. § Centrifugal.

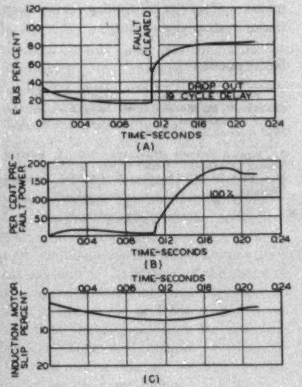


Fig. 5. Performance curves of motor 4, second study

A—Bus voltage
B—Per-cent prefault power
C—Motor slip

Second Study

A second transient study was made to determine the effect of the following major changes in the system:

1. Providing higher torque characteristics for motors *4* and *6*.
2. Adding time delay dropout characteristics for the contactors of motors *4*, *11*, and *14*.
3. Lowered dropout voltage of all 2,300-volt contactors.

Study A in Table III was made as shown in Fig. 1 and study B was made on the assumption that generator *2* was not available for service. Inspecting the results in Table III shows that system stability is essentially assured under both assumed conditions for the revised system. The studies were discontinued when the majority of the motors that had been dropped off because of undervoltage were brought back on the line and were accelerating towards normal speed.

For condition A, contactors for five motors dropped out, motor *5* being the only 2.3-kv motor to do so. The bus of motor *4* (2.3-kv) barely reached dropout value but the motor was not disconnected since its contactor had a time delay on dropout of approximately 0.317 second (19 cycles). Motors *6* and *8* (2.3-kv) were supplied through circuit breakers and were not dropped. The fault was assumed to be cleared in 0.11 second and by 0.22 second (time duration of study) three motors had been reconnected and were accelerating towards normal speed (motors *9*, *10*, and *13*). Motor *5* would have been connected after two more time increments (0.04 second). Motor *7* was disconnected because of the clearing of the fault and, since it would not be immediately reconnected because of relay action, its performance was not calculated.

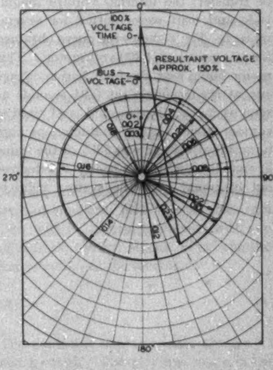


Fig. 6. Residual voltage drift curve for typical slow-speed motor near fault. Contactor dropout at 0.03 second and pickup at 0.23 second

Automatic transfer is controlled by a time delay undervoltage relay as well as a bus residual voltage relay set at 25%. All of the motors not dropped were accelerating rapidly. Three motors (*16*, *17*, and *18*) had assumed normal speed during the time duration of the study.

For condition B, contactors for seven motors dropped out. As in condition A, motor *5* was the only 2.3-kv motor to drop out. The fault was assumed to be cleared in 0.11 second and by 0.3 second (time duration of study); four motors had been reconnected and were accelerating towards normal speed. Motor *9* would be reconnected before the next step and motor *5* after the next step. Motor *7*, as in condition A, was again out of service. All motors not dropped were approaching normal speed, however, not as rapidly as in condition A. Motors *16*, *17*, and *18* had again assumed normal speed.

Marked improvement in the performance of motor *4* may be seen by comparing Fig. 5 with Fig. 3. By holding the contactor in through the use of 19 cycles' delay, the motor begins to reaccelerate as soon as the fault is cleared. Fig. 5 promises rapid recovery for this motor while Fig. 3 shows a definite tendency toward instability.

Residual Voltage Decay

Because of considerable interest in recent years in the magnitude and phase angles of motor residual voltages for circuits with rapid restoration of voltage, or transfer, the internal voltage magnitude and drift were plotted for several motors. Fig. 6 shows the curve obtained for an equivalent 2,300-volt motor. This motor being near the fault had a voltage rise at its terminals (at 0.04 second) as soon as it was disconnected. At time of contact touch, upon contactor pickup, its residual voltage was 52% of the value at time 0− and it was nearly 180 degrees out of phase with the system voltage. This gives an indication as to the possibility of damage due to transient torques and high reaccelerating current. Obviously those machines which drop off the line have their internal voltages pulled down, due to the proximity of the fault, and these residual voltages probably seldom exceed 50% upon instant of contact touch and would normally be lower.

By modifying the rectifier circuit for the d-c magnets of the 2,300-volt motor contactors, it is possible to delay dropout of these contactors for time intervals up to about 1/3 second. Thus, during a severe voltage dip lasting only about 1/10 second as assumed in these studies, the 2,300-volt motors which represent about 80% of the load could be kept connected to the system. This would overcome the contactor pickup time delay interval before motors would start reaccelerating their loads. And it would also overcome the hazard of reconnecting a motor with a high residual voltage when nearly 180 degrees out of phase with the system voltage.

Owing to the inherent design characteristics of induction motors, large motors are more likely to be damaged because of reclosure against high residual voltage. Therefore, the lower dropout voltage and the time delay dropout features were designed into the 2,300-volt contactors. Both of these features were accomplished with little additional cost and both tended to decrease the residual voltage problem on the large motors.

Ground-fault relay contacts can be connected in the contactor magnet control circuit to cause high-speed opening of the contactor if a ground fault should occur in a motor circuit. Since most faults will be of this nature, the protection afforded by ground-fault relaying of each 2,300-volt motor circuit is not sacrificed by the time delay dropout characteristic. On the other hand, this time delay characteristic and other features of the system design enables motors to ride through severe faults that might occur at rare instances and thus keep critical process operations in a refinery in continuous service. The entire system was designed with this end in view and the transient stability and voltage studies confirm that a high degree of reliability should be attained.

Conclusions

Studies testing the sufficiency in overall design of an industrial power distribution system are valuable in pointing out economies as well as deficiencies in design. This paper covers one method of attacking the transient stability problem in an oil refinery. It also shows how relatively inexpensive modifications can assure a large measure of service continuity against even major system disturbances.

References

1. TRANSIENT-STABILITY STUDIES—I, **D. F. Shankle, C. M. Murphy, R. W. Long, E. L. Harder.** *AIEE Transactions*, vol. 73, pt. III-B, 1954 (Feb. 1955 section), pp. 1563–80.

2. POWER SYSTEM STABILITY, VOL. I, **Edward Wilson Kimbark.** John Wiley & Sons, Inc., New York, N. Y., Chapman & Hall, Ltd., London, England, 1953.

A New Method for Compensating Contactor Servomechanisms

J. N. EMBLER
ASSOCIATE MEMBER AIEE

C. H. WEAVER
MEMBER AIEE

WHEN a servomechanism contains an on-off relay or contactor as a controlling device, it is called a contactor servomechanism. Among the great advantages of this type of servomechanism are its simplicity and low cost.[1] One of the great disadvantages, however, is that it falls in the category of nonlinear systems. Unfortunately, the state of the art at present is such that these systems do not lend themselves to analysis as readily as do conventional linear ones. Fig. 1 shows a diagram of a typical contactor servomechanism. Fig. 2 shows the block diagram of this same system.

Because of the fact that contactor servomechanisms are nonlinear, the analysis and synthesis of such systems have not been very successful until recent years. In fact, one of the most significant contributions in this field was the work of Kochenburger,[2] published in 1950. In his Doctor of Philosophy thesis at Massachusetts Institute of Technology, Kochenburger introduced the describing-function method of analyzing and synthesizing contactor servomechanisms. In this method, a linear describing function is used to approximate the nonlinear portion of the transfer function of the contactor. Of course, certain assumptions are made, and justifications of these assumptions are presented. It is beyond the scope of this paper to delve into such assumptions and justifications, and the reader is referred to the original work for complete details. However, certain of Kochenburger's results that are pertinent to this paper are discussed herein.

It is the purpose of this paper to develop a method of studying the relative stability of a contactor servomechanism that is more readily applied than that of Kochenburger. Whereas the latter method involves a certain amount of trial and error, this method does not; it entails construction, in the s plane, of root loci that are functions of amplitude rather than gain.

The construction of these root loci is greatly facilitated by the use of curvilinear squares, which make possible the transition from the $G(s)$-plane analysis to the s-plane analysis. Briefly, curvilinear squares may be described as follows: if, in the s-plane, the loci of s form a grid of squares, then their corresponding loci in the $G(s)$-plane will form a system of curves intersecting at right angles. In 1952, Russell and Weaver presented a method of synthesis of linear servomechanisms using such curvilinear squares to secure optimum location in the s-plane of the roots of the characteristic equation of the closed-loop transfer function.[3] These techniques will be drawn upon in constructing the root loci plots.

It is generally assumed in contactor-servomechanism theory that the contactors are instantaneous; i.e., there is no time lag between energization of relay and closing of contacts. The hysteresis effect, however, is not negligible. Since pickup current is appreciably greater than drop-out current on most relays practical for this application, this effect must be considered. Fig. 3 shows a contactor characteristic having an inactive zone Δ and a hysteresis zone h. It is this type of contactor which is discussed in this paper.

The first section gives a summary of the Kochenburger method of analyzing contactor servomechanisms. The second section summarizes the curvilinear square techniques as used by Russell and Weaver. The third section presents a study of contactor servomechanisms using root loci obtained from curvilinear square plots of Kochenburger's "linearized" systems. The fourth section summarizes the conclusions obtained from this study.

Summary of Kochenburger Method

As stated earlier, the Kochenburger method involves the use of a linear describing function which approximates the response of the nonlinear contactor. This approximation is accomplished by assuming a sinusoidal control-signal (contactor input) shape, analyzing the correction signal (contactor output) by the use of Fourier's series, and justifying the use of only the first harmonic of the output signal. Once this is done, the contactor response is linearized, and analysis greatly simplified. Kochenburger presents data to show that harmonics other than the first in general contribute so little to the contactor output that, even when they are neglected, the results usually fall within the range of engineering accuracy.

The use of this method results in the following relations between contactor output and input:

$$g_{D_1}=\frac{4\sin\beta\underline{/-\alpha}}{\pi|C_1|} \tag{1}$$

where

g_{D_1} = contactor describing function, output to input
C_1 = magnitude of contactor input, $C_1\cos\omega t$
β = angle at which corrective signal starts
α = angle of phase lag between contactor output and input

$$\beta=1/2\left[\cos^{-1}\left(\frac{\Delta-h}{2|C_1|}\right)+\cos^{-1}\left(\frac{\Delta+h}{2|C_1|}\right)\right] \tag{2}$$

$$\alpha=1/2\left[\cos^{-1}\left(\frac{\Delta-h}{2|C_1|}\right)-\cos^{-1}\left(\frac{\Delta+h}{2|C_1|}\right)\right] \tag{3}$$

Notice that the contactor describing function varies with the amplitude of the correction signal but not with frequency.

Fig. 4 shows a plot in the $G^{-1}(s)$-plane of the transfer function of the system shown in Fig. 1 except that no compensating network is included. Notice that there are two curves: one is the familiar $G^{-1}(s)$ plot of the linear portion of the system; the other is the contactor describing-function plot which does not vary with frequency, but rather with amplitude of control signal. The transfer function $G(s)$ of the servomotor is

$$G(s)=\frac{1}{s(1+s)} \tag{4}$$

The contactor characteristics are $\Delta=0.1$ and $h=0.02$. The plot was made in the inverse $G(s)$-plane because the desired results were more readily obtained from this plot than from the $G(s)$-plane plot. As Kochenburger points out, that portion of the contactor describing function that lies to the right of the $G^{-1}(s)$ locus, as the latter is traversed in a direction of increasing frequency, represents a region of instability. The point where the two loci intersect represents a point of stable oscillations. The region where the contactor locus lies to the left of the $G^{-1}(s)$ plot is a region of stability.

In this particular case, the system will always oscillate at the point of intersection of the two curves as a result of any disturbance great enough to actuate

---

Paper 57-393, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Southern District Meeting, Jackson, Miss., April 3-5, 1957. Manuscript submitted January 7, 1957; made available for printing March 22, 1957.

J. N. Embler is with the Aluminum Company of America, Alcoa, Tenn., and C. H. Weaver is with the University of Tennessee, Knoxville, Tenn.

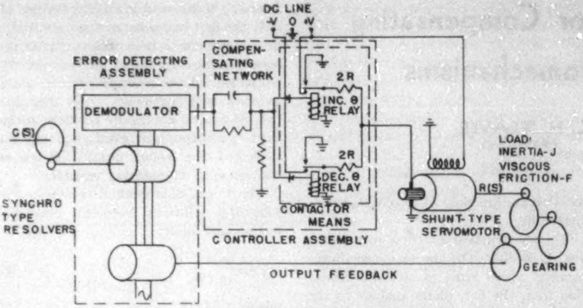


Fig. 1. Diagram of typical contactor servomechanism[9]

the contactor. If a small actuating signal is encountered, the contactor will operate in the unstable zone; thus oscillations will increase, giving rise to larger correction signals. This drives the contactor toward the intersection of the two loci. If a suitably large actuating signal is encountered, the contactor will operate in the stable region. In this case, oscillations decrease, correction signals become smaller, and the contactor again moves toward the intersection of the two loci. Thus stable oscillations result for any signal strong enough to actuate the contactor.

Fig. 5 is a plot of the same $G^{-1}(s)$ and the same contactor except that the inactive zone was increased. Here the entire range of the contactor is in the stable zone.

Notice at this point one significant difference between contactor and conventional servomechanisms. Consider the closed-loop transfer function of a conventional servomechanism

$$\frac{C(s)}{R(s)}=\frac{G(s)}{1+G(s)} \tag{5}$$

When

$$1+G(s_o)=0 \tag{6}$$

or, vectorially, when

$$\frac{1}{G(s_o)}=\underline{/180\text{ degrees}} \tag{7}$$

the expression $C(s)/R(s)$ has a pole at $s=s_o$. Furthermore, if the open-loop transfer function $G(s)$ is plotted in the $KG^{-1}(s)$-plane for one value of gain, the effect of gain change may be observed by moving the $-1+j0$ point along the real axis. As a matter of fact, the negative real axis is the root locus of the system with respect to curvilinear squares drawn in the $KG^{-1}(s)$-plane, as the gain K is changed. Now, consider the closed-loop transfer function of a contactor servomechanism

$$\frac{C(s)}{R(s)}=\frac{G(s)g_{D_1}}{1+G(s)g_{D_1}} \tag{8}$$

where the poles of $C(s)/R(s)$ now occur when

$$1+G(s)g_{D_1}=0 \tag{9}$$

or when

$$G(s)g_{D_1}=-1 \tag{10}$$

or when

$$G^{-1}(s)=-g_{D_1} \tag{11}$$

where

$$G^{-1}(s)=\frac{1}{G(s)} \tag{12}$$

From equation 11 it may be seen that if the open-loop transfer function of this system is plotted, the values of s for which $1+KG(s)g_{D_1}=0$, read on the curvilinear square plot, lie on the negative of the contactor describing-function locus and not on the negative real axis as the gain K takes on various values. Unlike the case of the conventional servomechanism, this shows that for the contactor servomechanism the curvilinear square root locus in the $G^{-1}(s)$-plane is the negative of the contactor describing-function locus.

Kochenburger demonstrates a graphic method (Fig. 5) of determining relative stability of such a system. Kochenburger shows that M_p, at any operating point Q on the contactor curve, is equal to the length of vector OQ divided by the length of the minimum vector QP, where point P is a point on the $G^{-1}(s)$ locus. Of course, the minimum QP will occur where $G^{-1}(s)$ is tangent to a circle of radius QP drawn from point Q. The point of tangency determines point P, and the frequency at which it occurs is the critical frequency of the system. It is readily seen that this method is an extension of the method of Brown and Campbell[8] for conventional servomechanisms. The principal difference between the two is that for the conventional system the center of the circle lies on the negative real axis.

In the system of Fig. 5, M_p, as Kochenburger defines it, approaches 4.5 near cutoff. Experience has shown that an M_p approachig 2.0 near cutoff, and about 1.3 for larger control-signal amplitudes will give satisfactory relative stability.

Fig. 6 is the plot of a system having the same $G(s)$ as that of Fig. 5, but a different contactor with a much smaller hysteresis zone. The inactive zone was made smaller too, because of accuracy considerations. This plot reveals a zone of instability like the system of Fig. 4. This system will, like that of Fig. 4, oscillate continuously as a result of any signal strong enough to actuate the contactor.

Fig. 7 involves exactly the same system as Fig. 6 with the addition of a simple phase-lead compensating network of the type shown in Fig. 8. This shifts the $G^{-1}(s)$ locus to the right, as shown in

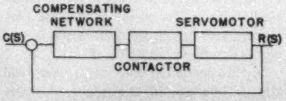


Fig. 2. Block diagram of contactor servomechanism

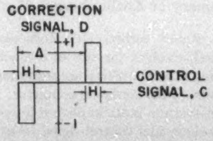


Fig. 3. Contactor characteristic showing inactive zone Δ, and hysteresis zone h

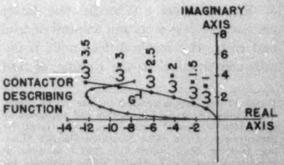


Fig. 4. Polar plot of G⁻¹ locus and contactor describing function locus for contactor servomechanism

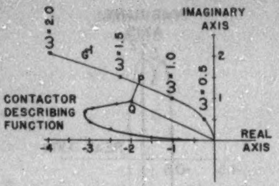


Fig. 5. Polar plot of G^{-1} locus and contactor describing function locus for contactor servomechanism

Fig. 7. By trial and error, Kochenburger arrived at the proper value of constants in the compensating network to give improved stability characteristics. In this case $M_p = 1.76$ at cutoff.

The third section of the paper presents a method for calculating the compensating network required for a given stability requirement rather than finding it by trial and error.

Summary of Curvilinear Square Techniques of Russell and Weaver

The use of curvilinear square plots as an aid to the analysis and synthesis of servomechanisms is not new. It was introduced in 1945 by Profos,[6] and has been used by various people in the field of servomechanisms since that time. Russell and Weaver used this method as a tool for predicting the location in the s-plane of the roots of the characteristic equation of the closed-loop transfer function. They also developed a method of compensating by which the amount of phase shift and/or attenuation required to shift such roots to more desirable locations could be predicted, and networks designed to give such shifts. These results are drawn upon for aid in the construction of root loci plots for the purpose of compensating contactor servomechanisms.

If, in the s-plane, the loci of s form a network of squares, then by conformal transformation to the $G(s)$-plane these loci may be represented by curves of constant ω and σ intersecting at right angles. The areas enclosed by this network of curves are called curvilinear squares. Fig. 9 shows a typical example. This applies equally well to the inverse $G(s)$-plane, and it is that plane that is used throughout this paper. It also applies to the Lm-angle plane, though the configuration will not be curvilinear squares.

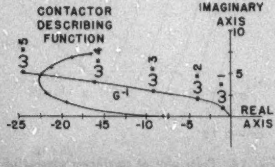


Fig. 6. Polar plot of G^{-1} locus and contactor describing function locus for contactor servomechanism

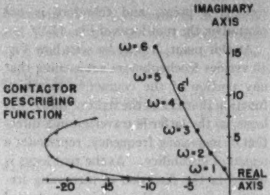


Fig. 7. Polar plot of compensated G^{-1} locus and contactor describing function locus for contactor servomechanism

Compensating networks can be designed to change root location a predetermined amount by changing either the phase angle or the magnitude, or both, of the $G(s)$ locus at a specific value of s. In compensating their systems, Russell and Weaver designed their compensating networks so that all roots would be shifted to the left of certain arbitrary lines in the s-plane, these lines being determined by the desired speed or response of the systems. Actually, the location in the s-plane of the roots of the characteristic equation of the transfer function of a given closed-loop system is the factor that determines the transient response of that system. It is also true that, in general, the farther these roots are to the left of the imaginary axis in the s-plane, the more relative stability the system will exhibit. When the designer knows what the speed of response and the relative stability of a system must be, he knows where the roots of the characteristic equation must lie in the s-plane. The techniques of Russell and Weaver are excellent means for designing compensating networks to place these roots in the desires location in the s-plane.

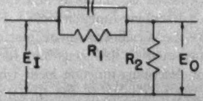


Fig. 8. Phase-lead compensating network

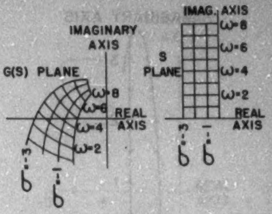


Fig. 9. Conformal transformation from s-plane to $G(s)$-plane

The next section shows a method of compensating a contactor servomechanism by root relocation in the s-plane.

A New Method for Compensating Contactor Servomechanisms

Two methods of analyzing servomechanisms have been reviewed: the Kochenburger method of approximating the nonlinear response of a contactor servomechanism by introducing a linear describing function to represent the contactor portion; and the Russell-Weaver curvilinear square technique for locating and shifting in the s-plane the roots of the characteristic equation of a closed-loop system. In this section, the two procedures are combined for the purpose of devising an improved method of analysis and synthesis of contactor servomechanisms. Curvilinear square plots are used to facilitate construction in the s-plane of the root loci plots of the characteristic equation. These root loci plots are analyzed from the standpoint of stability and compensation, and used to secure more favorable root locations.

As mentioned in the first section, Fig. 4 shows a plot in the $G^{-1}(s)$-plane of the transfer function of a contactor servomechanism. Fig. 10 shows the same curves with a curvilinear square plot of the frequency variant function superimposed.

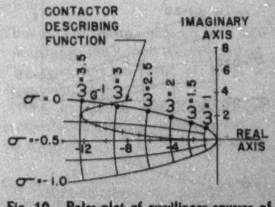


Fig. 10. Polar plot of curvilinear squares of G^{-1} locus and also contactor describing function locus for contactor servomechanism of Fig. 4

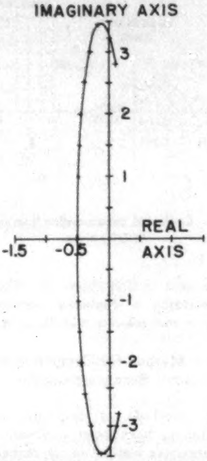


Fig. 11. Root locus in the s-plane of transfer function of contactor servomechanism of Fig. 10

Fig. 11 is the root locus plot in the s-plane of the complex roots of the characteristic equation plotted with amplitude of control signal as a parameter. This was obtained by applying the results of equation 11. As was pointed out in the first section, the root locus in the $G^{-1}(s)$-plane is the negative of the contactor describing function taken with respect to curvilinear squares. Since complex roots always occur in conjugate pairs, this curve is the locus of half the roots. The other half lie on a curve which is the conjugate of the contactor describing function shown in Fig. 10. Notice that on the curvilinear square plot the $\omega=0$ contour is a point, and therefore is not plotted on the root locus of Fig. 11.

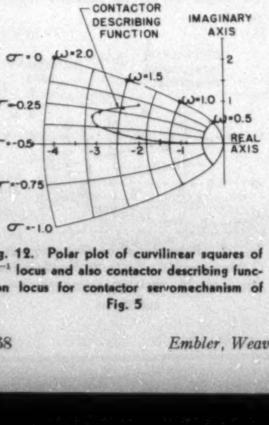


Fig. 12. Polar plot of curvilinear squares of G^{-1} locus and also contactor describing function locus for contactor servomechanism of Fig. 5

At this point, it may be seen how Fig. 10 verifies Kochenburger's reasoning that any portion of the contactor describing function that lies to the right of the $G^{-1}(s)$ locus, as the latter is traversed in a direction of increasing frequency, represents a region of instability. As the root locus is traced in the direction of increasing frequency, it crosses the imaginary axis and enters the right half of the s-plane, thus giving rise to instability.

Fig. 5 shows a plot of exactly the same system as that of Fig. 4 except that the inactive zone of the contactor was increased. Fig. 12 shows the same curves as Fig. 5, with a curvilinear square plot superimposed. Fig. 13 is the root locus plot of this system. Here again, Kochenburger's reasoning that the entire amplitude range of the contactor is within a stable operating region is confirmed. As shown by the root locus, there are no roots on the imaginary axis or in the right half of the s-plane.

Fig. 6 represents the $G^{-1}(s)$ locus and contactor locus of a different contactor servomechanism having both a smaller inactive zone and a smaller hysteresis zone. Fig. 14 shows these characteristics with curvilinear squares added. Fig. 15 is the root locus of this system, and reveals roots in the right half of the s-plane, giving rise to instability.

Fig. 7 shows the $G^{-1}(s)$ locus of this system after the addition of a phase-lead compensating network. Fig. 16 is Fig. 7 with curvilinear squares added. Fig. 17 is the root locus of the compensated system, showing the roots shifted well to the left in the s-plane.

Up to this point, curvilinear squares have been used merely as an aid to studying the results of Kochenburger. Now it will be shown how they may be used directly as an aid to the compensation of contactor servomechanisms. First, they will be used to determine the location, in the s-plane, of the roots of the characteristic equation of Kochenburger's compensated system. Second, as a check, they will be used to duplicate his results without the use of trial-and-error methods.

The location of roots in the s plane is entirely a graphical process and is shown on Figs. 7, 16, and 17. Fig. 7 shows the compensated $G^{-1}(s)$ locus and the contactor describing-function locus. Fig. 16 shows the curvilinear square plot of the compensated $G^{-1}(s)$ locus and the contactor describing-function locus which is also the root locus in the $G^{-1}(s)$-plane. Fig. 17, the root locus of the system in the s-plane, is obtained by plotting the root locus of Fig. 16 with respect to the curvilinear squares.

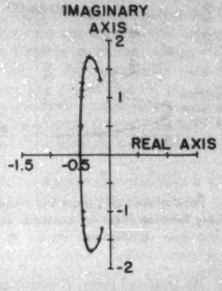


Fig. 13. Root locus in s-plane of transfer function of contactor servomechanism of Fig. 12

Before proceeding with the calculation of the compensating network, it is necessary to know two details about the system in order to apply the Russell-Weaver techniques. First, since this process is one of root shifting, specific roots must be chosen. This means that some particular operating point on the contactor describing function must be chosen. For demonstration purposes the cut-off point is used. Cutoff may be defined as the point on the contactor describing function that represents the minimum signal that will operate the contactor. That is to say, the contactor will not respond to any input signal of less magnitude than the one at cutoff. It should be noticed that this is also the point of maximum M_p as Kochenburger defines it. Second, it is necessary to

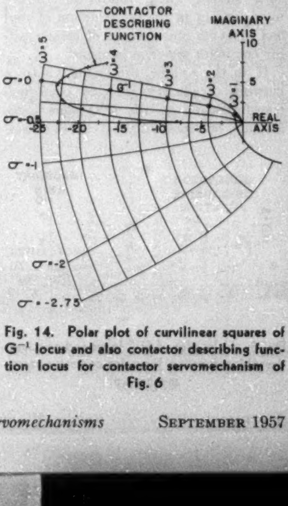


Fig. 14. Polar plot of curvilinear squares of G^{-1} locus and also contactor describing function locus for contactor servomechanism of Fig. 6

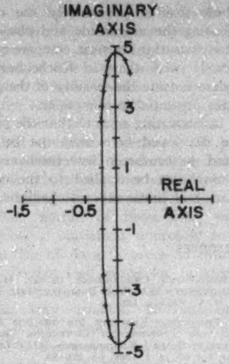


Fig. 15. Root locus in the s-plane of transfer function of contactor servomechanism of Fig. 14

know the phase angle between the location of the roots before and after shifting. (This angle is designated as θ in the Russell-Weaver equation for calculating compensating networks.) This angle may be determined from curvilinear square plots. In this case, the root locus of Fig. 17 shows that when the compensated system is operating at cutoff, the root of the characteristic equation is located at $-\sigma=-2.75$ and $\omega=4.75$. It is seen by referring to Fig. 14, the curvilinear square plot of this system before compensation, that the phase at this point is 230.3 degrees. Fig. 14 shows the cut-off point before compensation also, and the phase at this point is seen to be 156.0 degrees. Therefore θ, the angle of required phase shift, is $230.3-156.0=74.3$ degrees.

The compensating network may now be calculated. The transfer function of the compensating network shown on Fig. 8 according to the Russell-Weaver method is

$$\frac{E_0(s)}{E_i(s)}=G_2(s)=\frac{\frac{1}{\alpha_l\,\tau_l}+s}{\frac{1}{\tau_l}+s}\tag{13}$$

where

$$\tau_l=\frac{R_1R_2C}{R_1+R_2}\tag{14}$$

and

$$\alpha_l=\frac{R_1+R_2}{R_2}\tag{15}$$

In practice, values of α_l between 5 and 20 have been found most satisfactory. Equation 12 of Appendix II of Russell and Weaver[7] may be used to determine τ_l of the compensating network. Thus

$$\alpha_lT^2(1+\tan^2\psi_k)-T\left[\alpha_l+1+\frac{\tan\psi_k(\alpha_l-1)}{\tan\theta}\right]+1=0\tag{16}$$

where

$T=\tau_l\alpha_k$

$\psi_k=\tan^{-1}\frac{\beta_k}{\alpha_k}$

θ = angle of required phase shift

$-\alpha_k+j\beta_k$ = point in s-plane to which root is being relocated

$\tan\psi_k=\frac{4.75}{2.75}=1.73$

$\tan\theta=\tan 74.3=3.56$

$\alpha_l=10$ arbitrarily

$T=\alpha_k\tau_l=2.75\ \tau_l$

Applying equation 16 gives $T=0.382'$ from which $\tau_l=0.139$ and $R_1C=1.39$. The transfer function of the compensating network may now be written:

$$G_2(s)=\frac{\frac{1}{(10)(0.139)}+s}{\frac{1}{(0.139)}+s}=\frac{0.178+s}{7.18+s}\tag{17}$$

Since this transfer function is not in the same form as the Kochenburger transfer function for the compensating network, the two may not be directly compared. However, as a check, both networks will be evaluated at the point in question and the results compared.

For the Russell-Weaver network

$$G_2(s)=\frac{\frac{1}{(10)(0.139)}+s}{\frac{1}{(0.139)}+s}\tag{18}$$

$$G_2(s)=\frac{0.718+s}{7.18+s}\tag{19}$$

$$G_2^{-1}(s)=\frac{7.18+s}{0.718+s}\tag{20}$$

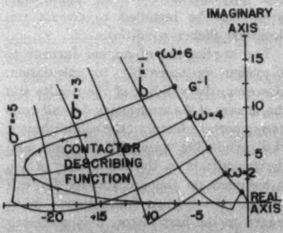


Fig. 16. Polar plot of curvilinear squares of G^{-1} locus and also contactor describing function locus for compensated contactor servomechanism of Fig. 7

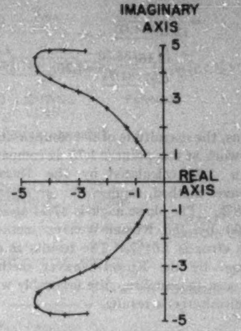


Fig. 17. Root locus in the s-plane of transfer function of contactor servomechanism of Fig. 16

Evaluated at $s=j4.75$,

$$G_2^{-1}(j4.75)=\frac{7.18+j4.75}{0.718+j4.75}=\frac{8.61\underline{/33.50}}{4.80\underline{/81.40}}=1.79\underline{/-47.90}\tag{21}$$

For the Kochenburger network[8]

$$G_2(s)=\frac{E_0s}{E_i(s)}=\frac{R_2}{R_0+R_1+R_2}\,\frac{(1+s\tau_a)}{(1+s\tau_a/\alpha)}\tag{22}$$

where

$$\tau_a=R_1C_1\tag{23}$$

$$\alpha=\frac{R_0+R_1+R_2}{R_0+R_2}\tag{24}$$

R_0 = internal resistance of network preceding compensating network. (If this resistance is neglected, the transfer function is identical to that of Russell and Weaver.)

Kochenburger used the following values in his compensating network:

$$\begin{aligned}R_0&=4{,}000\\R_1&=1.45\times10^6\\R_2&=157{,}000\\C_1&=1.0\times10^{-6}\\\alpha&=10\end{aligned}\tag{25}$$

Thus $R_1C_1=1.45$ as compared with 1.39 calculated by the Russell-Weaver method. This represents an error of 4.0% in determining R_1C_1.

Now, evaluating the Kochenburger compensating network yields the following equations

$$G_2(s)=\frac{1}{10}\,\frac{(1+1.45s)}{(1+0.145s)}\tag{26}$$

Evaluated at $s=j4.75$

$$G_2(j4.75)=\frac{1}{10}\,\frac{(1+1.45\times j4.75)}{(1+0.145\times j4.75)}=\frac{1+j6.89}{10+j6.89}$$

$$=\frac{6.97\underline{/81.74}}{12.15\underline{/34.50}} \qquad (27)$$

$$G_1^{-1}(j4.75)=\frac{12.15\underline{/34.50}}{6.97\underline{/81.74}}=1.75\underline{/-47.14} \qquad (28)$$

Thus, the magnitude of the compensating network at this point is 1.75 as compared with 1.79 calculated by the Russell-Weaver method, representing an error of 2.28%. The phase angle is 47.14 against 47.90 by the Russell-Weaver method; this error is 1.61%. The results as obtained by the Russell-Weaver method are seen to compare quite favorably with Kochenburger's results.

Conclusions

By trial-and-error procedures, Kochenburger demonstrated a method of compensating contactor servomechanisms. Recognizing that trial-and-error procedures are at best tiresome, tedious processes, a method of determining such compensation directly has been developed and presented in this paper. The basis for this method is the placement of the roots of the characteristic equation of the system in suitable locations in the s-plane by means of a combination of root locus and curvilinear square techniques. In general the desired location of such roots can be determined whenever the desired stability and speed of response of the system to be compensated are known.

The method developed in this paper is particularly useful because it is impossible to use logarithmic (Bode) plots with the describing-function type of synthesis. This means that Kochenburger's cut-and-try methods must be carried out on the polar $G(s)$ plot, and hence are quite difficult.

In his paper, Kochenburger applied his trial and error method to the determination of a lead network for compensation of a given contactor system. To check the method developed in this paper, a similar network was determined. The results of the two calculations compare as follows:

	Kochenburger	Calculated	Error, Per Cent
Time constant, R_1C_1	1.45	1.39	4.14
Magnitude of compensating network at $s=4.75$	1.75	1.79	2.28
Phase of compensating network at $s=4.75$	47.14°	47.90°	1.61

These results, particularly the ones concerning the magnitude and phase of the compensating network, compare quite favorably with those of Kochenburger, and demonstrate the feasility of the procedures presented in this paper.

It is interesting to note that the technique developed here need not be restricted to contactor servomechanisms, but may also be applied to the compensation of linear servomechanisms.

References

1. Engineering Cybernetics (book), H. S. Tsein. McGraw-Hill Book Company, Inc., New York, N. Y., 1954, p. 136.

2. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, **Ralph J. Kochenburger**. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

3. Synthesis of Closed-Loop Systems Using Curvilinear Squares to Predict Root Location, D. W. Russell, C. H. Weaver. *Ibid.*, vol. 71, pt. II, 1952, pp. 95–104.

4. **Kochenburger**, *loc. cit.*, p. 278.

5. Principles of Servomechanisms (book), Gordon S. Brown, Donald P. Campbell. John Wiley & Sons, Inc., New York, N. Y., 1948, p. 188.

6. A New Method for the Treatment of Regulation Problems, P. Profos. *Sulzer Technical Review*, New York, N. Y., vol. 2, 1945, pp. 1–5.

7. **Russell, Weaver**, *loc. cit.*, p. 104.

8. **Kochenburger**, *loc. cit.*, p. 272.

9. **Kochenburger**, *loc. cit.*, p. 271.

Vo
Pt.
19

Survey of Arc-Furnace Installations on Power Systems and Resulting Lamp Flicker

AIEE SUBCOMMITTEE REPORT

THE ELECTRIC arc furnace is usually considered a desirable load because of its good load factor and its high power factor. However the violent load variations, predominantly single-phase, which occur particularly during melt-down, make it a difficult load to supply because of the objectionable voltage fluctuations and lamp flicker which can result. The magnitude of these load fluctuations compared to the average furnace load is much greater than that existing for most large industrial loads. Designing a power system for supplying a large arc furnace requires much greater attention to flicker considerations than other types of loads and may actually require that special equipment be installed to correct undesirable flicker conditions.

The application of an arc furnace to a system is amenable to calculation. However, the validity of the results and the success of the application depend upon two important factors: the establishment of a permissible voltage flicker limit, and the magnitude of the load swing which is assumed on the arc furnace for calculating the voltage fluctuations. With regard to the first factor, it is necessary to correlate the magnitude and frequency of voltage changes with the change in lumen output of a lamp as it is discerned by the human eye. The varying rates of furnace swings make it difficult to select one frequency which will characterize the annoyance to the average viewer. Hence, partially because of the uncertainty regarding effective swing frequency, a difference of opinion exists in the industry as to the magnitude of voltage fluctuations due to arc furnaces which can be allowed.

With regard to the second factor, the load swings produced by an arc furnace are difficult to predict. Extensive basic data have been obtained on existing furnace installations by means of oscillographs or recording meters. These

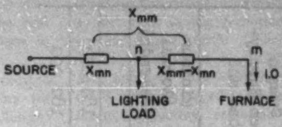
---

Paper 57-9, recommended by the AIEE System Engineering Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 22, 1956; made available for printing October 31, 1956.

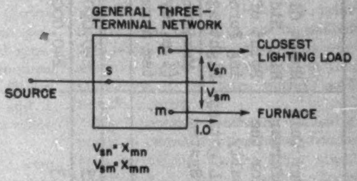
This survey was conducted by the Working Group on Arc Furnaces, under the direction of the System Planning Subcommittee of the Committee on System Engineering. The membership of the Working Group was as follows: B. L. Lloyd, *chairman;* R. W. Baker, D. H. Cameron, L. K. Del'Homme, C. H. Didriksen, A. P. Fugill, A. C. Hartranft, H. O. Hodson, B. M. Jones, J. H. Kinghorn, R. F. Lawrence, L. F. Lischer, W. H. Osterle, L. R. Patterson, V. G. Rettig, T. W. Schroeder, G. L. Sidway, J. E. Treweek, M. C. Westrate.

data have been very valuable in the past in guiding arc furnace installations. However, the load swings produced by a furnace are dependent on a large number of factors which vary from time to time and from installation to installation. Some of these factors are the method of furnace operation, the grade of scrap steel used, transformer secondary lead length, transformer secondary voltage, billing considerations, etc. The random nature of these load swings requires a statistical approach to the collection of data in order to establish an effective load swing for furnaces of various sizes, and different combinations of the many variables. No single utility has a sufficient number of furnace installations of the various sizes or the time, equipment, and personnel required to accumulate this extensive mass of data. The situation was recognized by the AIEE Committee on System Engineering, and in 1953 the Working Group on Arc Furnaces was organized in an attempt to eliminate some of these uncertainties.

Fig. 1. Illustration of definition and significance of self-drop constant (X_{mm}) and mutual drop constant (X_{mn})



(A) SIMPLE RADIAL CIRCUIT



(B) GENERAL THREE TERMINAL NETWORK

The purpose of the Working Group was to determine whether future installations of arc furnaces involving flicker considerations could be based on past operating experiences and practices. The general procedure decided upon was to attempt some correlation of data by means of a survey among utilities which had existing installations of furnaces. The decision was made quite early not to attempt determination of furnace load swings. While these data would have been very valuable, accumulating them on a large number of furnaces would have been extremely difficult and costly. Instead, it was decided to develop a yardstick of the characteristics of the supply system and correlate these characteristics with the flicker performance which had been experienced on these existing furnaces.

The survey was conducted and an unpublished preliminary report of the results was made in 1955. The present paper gives the final results of the survey and the correlation which was obtained and suggests a method for determining the suitability of an arc furnace application to a power system.

Significance of Self- and Mutual Drop Constants

Many past efforts to correlate system requirements with furnace size have expressed the system requirement in terms of the 3-phase short-circuit capacity at the bus to which the furnace transformer is connected. This short-circuit capacity is the reciprocal of the equivalent system impedance from that bus to the equivalent internal voltage of the system generators. For a given load swing, it is apparent that the voltage drop due to this swing (or flicker voltage) on the furnace transformer bus is directly proportional to the magnitude of this equivalent system reactance. Therefore it is possible to obtain a correlation of furnace size and short-circuit capacity at the furnace transformer bus so long as the flicker voltage on that bus is the determining quantity.

However, only in the minority of cases is the closest lighting load connected directly to the furnace transformer bus. In the majority of cases this closest lighting load will be connected back at some intermediate point in the system. In fact the survey reveals that in some cases the flicker voltage at the closest lighting load was less than 20% of the flicker voltage which occurred at the furnace transformer bus. Consequently, unless the closest lighting load is connected at or very near the furnace transformer bus, the short-circuit capacity on that bus does not provide a suitable criterion of objectionable lamp flicker on the utility system.

The Working Group selected a different and more fundamental quantity in the attempt to correlate furnace size and system requirements. This quantity is designated "mutual drop constant" and is identified by the symbol X_{mn}, the definition of which can be most readily understood by referring to the simple radial system shown in Fig. 1(A).

The furnace transformer bus is designated as *m* and the location of the closest lighting load as *n*. With 1.0 per-unit reactive current (based on the furnace transformer name-plate kva) drawn from the system at point *m*, X_{mn} is defined as the voltage drop from the generator equivalent internal voltage to the point *n*. The quantity X_{mm} is defined as the voltage drop from the generator equivalent internal voltage to the point *m*, and this quantity is designated as the "self-drop constant." The quantity X_{mn} is the reciprocal of the short-circuit capacity at the furnace transformer bus. The difference between X_{mn} and X_{mm} as shown in Fig. 1(A) further illustrates the previous statement that the short-circuit capacity on the furnace transformer bus is not a good index of objectionable flicker voltage if a lighting load is not connected to that bus. For the simple radial circuit in Fig. 1(A), X_{mn} is the reciprocal of the short-circuit capacity at the lighting tap, but in more general networks, this is not the case.

For the purpose of explanation, the simple radial circuit of Fig. 1(A) was chosen merely for its simplicity and the clarity with which it illustrates the concept and the significance of the self- and mutual drop constants. Following the same procedures will permit determination of X_{mn} and X_{mm} for the general 3-terminal network as shown in Fig. 1(B).

The assumptions used in this survey for setting up the 3-terminal network for the various furnaces covered by the survey are as follows:

1. Voltage of generators and synchronous condensers are equal and in phase.
2. Shunt impedance branches are neglected.
3. Subtransient reactance is used for all synchronous machines.
4. The resistance in lines, transformers, and generators is neglected.

1 Arc Furnace Ident.	2 Furnace Transf. Kva Rating. (List each separately) (Rating refers to that on which impedance is given)	3 Method of Furnace Operation (Check one)		4 Type of Control Used on Furnace	5 X_{mn} in Per Unit on Transformer Rating. (System normal)	6 X_{mn} in Per Unit on Transformer Rating for the Electrically Closest Lighting Load. (System normal)	7 Short Circuit Kva on High Side, at Common Point of System Supplying Furnace Load and Lighting Load	
		Cold Melt	Refining				Maximum (i.e. system normal)	Minimum (Normal Weekends)
				COMPANY "J"				
a-1	3500		✓	Contact-Making Ammeter	0.057	0.013	240,000	180,000
a-2	3500		✓	Regulex	0.057	0.013	240,000	180,000

8 Corrective Measures which are Employed to Minimize Flicker. (Check one which applies)								9 Type of Lighting Load (Check One)	
Synchronous Condenser	Syn. Cond. & Buffer Reactor	Series Capacitor	Flicker Bus	Separate Radial Line	Capacitor in Series with Syn. Cond.	None	Other (Describe)	Industrial	Residential
✓				✓					✓
✓				✓					✓

10 Is Flicker Objectionable (Cust. Complaints) at Lighting Load?			11 If corrective Measures per Column 8 were Removed, What Would be X_{mn}?	12 For Column 11, Would the Flicker be Objectionable?	13 If answer to Column 10 is Yes, is the Objectionable Flicker Serious Enough to Warrant Relief Measures?	14 State any Condition of System Operation Under Which Flicker is Objectionable.	15 For Column 14, What is X_{mn} for this Condition?
Yes	Borderline	No					
	✓		.023	Yes		7500 kva or more gen.	.017
	✓		.023	Yes		Shutdown at nearby plant	.017

Fig. 2. Sample of form-sheet used in survey

The self- and mutual-drop constants can be determined analytically by network reduction or by use of a d-c or a-c calculating board.

The use of 1.0 per-unit current out of the network at point m for determining the drop constants must not be construed to mean that the Working Group assumes that the actual swing for the particular furnace is 1.0 per unit. This arbitrarily assumed current merely provides a convenient kva base for expressing the drop constants.

In locating point n, consideration should be given to the possibility of having a future lighting load closer to the furnace installation than exists at present.

Description of Survey

The representatives on the Working Group determined X_{mn} for the furnaces on their respective utility systems. They also supplied other pertinent data as requested on the survey form sheet. A sample of the survey form sheet is shown in Fig. 2, and the most significant data are shown in Table I. Each furnace is identified by a capital letter which designates the utility company; a lowercase letter which designates the particular furnace installation on the system; and in the case of multifurnace installations, an arabic numeral designates a particular furnace within an installation. The value of X_{mn} for normal system operation was requested in column 6 of Fig. 2 along with a statement as to whether the normal flicker condition was satisfactory, objectionable, or borderline. Information was requested as to the type of corrective measures employed to control flicker. If the corrective measures were removed, column 11 requested the new value of X_{mn}, and a statement was requested as to whether the flicker would then be objectionable. Column 15 requested X_{mn} for any other system condition at which flicker would become objectionable. All this information accumulated by the survey is summarized and tabulated in Table I.

The tabulated data received from the various utilities was usually accompanied by comments which were helpful in analyzing the data. These comments are included in the Appendix.

Analysis of Survey Data

The furnace installations reported in the survey consisted of from 1 to 13 furnaces per installation. The correlation of all these data is rather difficult because of the variation in the number and size of furnaces in the various installations. It was found that the best correlation obtained is that shown in Figs. 3 through 5.

Fig. 3 shows the value of X_{mn} plotted against furnace transformer rating for each of the single-furnace installations tabulated in the survey. The various symbols in Fig. 3 indicate whether the lamp flicker was judged to be objectionable, borderline, or satisfactory. Curves in Fig. 3 indicate the maximum permissible value of X_{mn} for assured satisfactory flicker conditions, the minimum X_{mn} above which objectionable flicker should definitely be expected, and a marginal area between these curves where the flicker is in a borderline region.

All the points shown in Fig. 3 were not obtained from installations where only one furnace was present. The survey listed a number of cases where one large furnace existed at an installation along with one or more smaller furnaces. For cases where the large furnace represented a major part of the installed furnace capacity and where this large furnace was 25% larger than the next size, the point for this largest furnace was plotted as a single furnace in Fig. 3. Ignoring the smaller furnaces at an installation permits use of a great bulk of survey data which would otherwise be useless. Previous industry literature estimates that the effect of adding a second furnace of the same size as an existing furnace will increase the load swings by 15% to 20%. Adding a second furnace which is substantially smaller than the first will hardly increase the swings at all. Hence it is believed that ignoring these smaller furnaces at an installation will not decrease the accuracy of the results.

Fig. 4 shows plots of X_{mn} against furnace size for installations with two furnaces of identical rating. Fig. 5 shows similar data for installations with three or more furnaces of the same rating. There were so few installations of more than three identical furnaces at a location that no separate curves were attempted beyond the 3-furnace limit. Instead, all installations with three or more identical furnaces were plotted in Fig. 5. This is permissible because there will be only a very small difference between X_{mn} required for three or for more than three furnaces. Again, as in Fig. 3, some of the points on Figs. 4 and 5 ignore the presence of some smaller furnaces at the installations plotted.

Some points on the multifurnace curves do not actually consist of identically sized furnaces in the numbers indicated by the curves. This is illustrated by the point in Fig. 5 plotted at 16,000 kva and X_{mn} of 0.0065. This installation actually consists of an 18,000-kva furnace and two 16,000-kva furnaces, and the flicker was judged to be satisfactory. Since satisfactory performance is obtained with these furnaces, the performance would certainly be satisfactory if the installation consisted of three 16,000-kva furnaces. Hence the point indicated on the curve can be plotted with complete assurance that satisfactory operation would be experienced at this installation if the three smaller sized furnaces existed.

Fig. 3 contains a sufficient number of points to permit drawing rather definite limit curves for furnaces below about 15,000 kva. Fig. 4 contains fewer points than does Fig. 3, but the limit curves will still be reasonably accurate. The limits for 2-furnace installations lie about 10% to 15% below the curves for single-furnace installations. The scarcity of points in Fig. 5 makes it impossible to draw accurate limit curves from the available data. Somewhat arbitrary ones are drawn at about 5% to 10% below the curves for 2-furnace installations.

The general shape of the limit curves is approximately that which would be expected by consideration of previous literature. The per-unit load swing experienced by a small furnace is greater than that experienced by a larger furnace. Consequently, with the same tolerable flicker voltage limit, the larger furnaces permit a higher per-unit value of X_{mn}. Note that X_{mn} is in per-unit on the furnace transformer kva base.

Cyclic Flicker in Large Furnaces

The limit curves in Figs. 3, 4, and 5 are shown as solid lines for furnaces from 1,000 to 15,000 kva, and are shown dashed over 15,000 kva. For the smaller furnaces below 15,000 kva, the resulting lamp flicker appears to occur at rates from one per minute to one per second, which is termed the "extremely frequent" range. There is some evidence that for furnaces larger than 15,000 kva, "cyclic-flicker" voltages at a rate greater than one per second will be superimposed on the extremely frequent flicker voltages. The tolerable level of flicker voltage in this cyclic range may be as low as one half to one fourth of the tolerable level in the extremely frequent range. Hence cyclic load swings can be smaller in magnitude than the extremely frequent swings, yet still be more objectionable.

The effect of this cyclic flicker is

shown by the 25,000-kva points in Fig. 4. Borderline and objectionable flicker are shown for values of X_{mn} which are far below the limit curves obtained by extrapolating data for the smaller furnaces. An indication of some lower limit is also shown by one of the 25,000-kva points in Fig. 5. However, Figs. 3, 4, and 5 show a number of installations from 15,000 to 25,000 kva where the data correlate quite well with the indicated limit curves. Hence for these large furnaces there is apparently something about the characteristics of the particular furnace which determines whether that furnace will produce objectionable swings at the cyclic rate.

There has been speculation about whether the large cyclic swings are due to the size of the furnace or to higher electrode voltage. Almost all of the furnaces covered in this survey have an electrode voltage in the order of 350 volts or lower. However, some of the newer large furnaces have much higher electrode voltages, and the resulting load swings may differ from those of the furnaces covered in the survey. The theory attributing cyclic swings to the higher electrode voltage is in part substantiated by the installations described in the preceding paragraph. The 25,000-kva furnaces in Fig. 4 are now operating with an electrode voltage of 510 volts, and those in Fig. 5 have a maximum transformer secondary tap of 445 volts. However, it should be pointed out that one company has reported cyclic swings, although not of objectionable magnitude, on 16,000-kva furnaces operating at the lower electrode voltages.

There are not a sufficient number of these large furnaces which have experienced trouble due to cyclic flicker to permit the working group to develop definite limits of X_{mn} for these large sizes. This will become an increasingly important problem in the future as ever larger furnaces are required for steelmaking; unfortunately the problem could not be resolved with the scope of this survey. Consequently all the limit curves in Figs. 3, 4, and 5 are shown dotted above 15,000 kva, and the working group cautions that cyclic flicker may exist above this rating and may require lower limits of X_{mn} than are indicated by the curves.

Table I. Summary of Results of Survey

Furnace Identification*	Furnace Transformer, Kva	Normal X_{mn}, Column 6†	X_{mn}, Column 11†	X_{mn}, Column 15†	Corrective Measures‡
A-a-1,2	20,000	0.0184			
a-3-5	12,000	0.0111			
b-1	12,000	0.0040			
b-2-6	12,000	0.0040			
b-7	10,000	0.0033			
b-8	10,000	0.0033			
b-9	7,500	0.0025			
b-10,11	4,500	0.0015			
b-12	3,000	0.0010			
c-1	7,500	0.0056			
c-2,3	10,000	0.0074			
c-4-6	20,000	0.0148			
c-7	15,000	0.0111			
c-8,9	12,000	0.0089			
d-1	7,500	0.0052	0.0121§		ST & SRL
d-2,3,5	3,000	0.0021	0.0049§		ST & SRL
d-4	6,000	0.0041	0.0097§		ST & SRL
B-a-1	7,500	0.0091	0.0153‖		SRL
a-2	4,000	0.0048	0.0082‖		SRL
b-1,2	1,500	0.0027			
c-1	1,200	0.0059			
c-2	1,200	0.0059			
d-1-4	7,500	0.0070			FB
e-1	2,000	0.0019	0.0037‖		ST & SRL
e-2	2,250	0.0021	0.0042‖		ST & SRL
e-3	5,000	0.0047	0.0094‖		ST & SRL
f	1,500	0.0014	0.0028		SRL
C-a	12,500	0.0069	0.0180‖		ST & SRL
b	4,000	0.0095§			
D-a-1	1,200	0.0064§			
a-2	1,500	0.0080§		0.020‖	
b-1	800	0.0037			
b-2	1,200	0.0055		0.025‖	
c-1	12,000	0.0117	0.0220‖		SRL
c-2,3	18,000	0.0175§	0.0340‖		SRL
d	2,500	0.0060			
e-1,2	15,000	0.0095	0.0190§		SRL
E-a-1	500	0.0006	0.0037		SRL
a-2	1,000	0.0013	0.0074‖		SRL
a-3	2,000	0.0026	0.0147‖		SRL
a-4	3,000	0.0039	0.0221‖	0.0091‖	SRL
b-1-6	3,000	0.0072	0.0165‖		SRL
b-7,8	1,000	0.0024	0.0055		SRL
c-1	18,000	0.0074			SC
	18,000	0.0094			SC
c-2,3	16,000	0.0065			SC
c-4	12,000	0.0049			SC
d-1	1,500	0.0052	0.0061		
d-2	2,500	0.0088§	0.0101‖		
F-a-1,2	7,500	0.0036			2 SRL
b	3,500	0.0110§			
c	3,500	0.0023			SRL
d-1-3	4,000	0.0021			
e	2,000	0.0044			
f	2,000	0.0121			
g-1-3	2,000	0.0118§			
h-1	1,800	0.0108			
h-2	2,000	0.0120			
i-1	3,000	0.0170			
i-2	1,500	0.0085			
j-1,2	7,500	0.0064	0.0153‖		SRL, SC & BR
k-1,2	4,000	0.0085§			
k-3	3,000	0.0064§			
k-4	2,500	0.0053			
l-1,2	4,000	0.0120			
l-3	3,000	0.0090			
l-4	2,500	0.0075			
m-1	12,000	0.0027			
m-2,3	15,000	0.0034			
n-1-6	12,000	0.0027			
n-7-9	20,000	0.0046			
o	1,600	0.0050			
p-1,2	25,000	0.0080§	0.020‖		SC & BR
p-3	25,000	0.0080§	0.020‖		SC & CS
p-4	15,000	0.0050	0.012‖		SC & CS
q-1-4	3,000	0.0042			
r-1,2	3,000	0.0058			
s	2,000	0.0044			
t	3,000	0.0157§			
G-a-1	1,500	0.0031	0.0058		SRL
a-2	4,000	0.0084	0.0155‖		SRL
b-1,2	6,000	0.0048			
b-3	12,000	0.0096			
b-4	18,000	0.0144			
c-1,2	10,000	0.0186	0.0302‖		
d-1-4	10,000	0.0018	0.0259‖		C
d-5	18,000	0.0032	0.0465‖		C
e-1	2,500	0.0066			
e-2	3,000	0.0079			
f-1	2,500	0.0062			
f-2	2,000	0.0050			
g-1	3,000	0.0100		0.018‖	
g-2	5,000	0.0166§		0.0301‖	
h	5,000	0.0159		0.0230‖	

Use of Subtransient Versus Transient Reactance

It was stated previously that the data tabulated in Table I were obtained with the use of subtransient reactance for all

Table I. (Cont.)

Furnace Identification*	Furnace Transformer, Kva	Normal X_{mn}, Column 6†	X_{mn}, Column 11†	X_{mn}, Column 15†	Corrective Measures‡
i	2,000	0.0068			
j-1,2	4,000	0.0076			
k-1,2	4,000	0.0046			
k-3	5,000	0.0058			
k-4	2,000	0.0023			
k-5	5,000	0.0058			
l-1	12,000	0.0033			
l-2,3	15,000	0.0042			
l-4	6,000	0.0017			
m	5,000	0.0014		0.0199‖	
n-1	2,000	0.0081			
n-2	1,200	0.0049			
o	2,000	0.0089		0.0204‖	
H-a	15,000	0.0086	0.0268‖		FB
b	1,800	0.0094	0.0276‖		SRL
c	3,300	0.0069	0.0220‖		SRL
I-a	25,000	0.0060	0.018‖		SC & BR
b	7,500	0.0044	0.016‖		SC & BR
J-a-1,2	3,500	0.013§	0.023‖	0.017‖	SRL & SC
K-a	25,000	0.029		0.057‖	FB
L-a	2,500	0.0097			
b	2,000	0.0062			
c	2,500	0.0014			SRL
d-1	10,000	0.0056			SRL
d-2	5,000	0.0028			SRL
e	2,250	0.0065			
f	2,500	0.0071			
g	1,500	0.012			
h-1	2,000	0.0057			
h-2	2,500	0.0071			
M-a	1,800	0.0016			
b	2,000	0.0031			
N-a-1	15,000	0.0238‖			
	15,000	0.0129§			
	15,000	0.0072			SRL
a-2	18,750	0.0162§			
	18,750	0.0090			SRL
a-3	25,000	0.0215§			
	25,000	0.0120			SRL
O-a	5,000	0.0065	0.0622‖	0.049‖	C
P-a-1,2	16,000	0.0043			SRL
a-3,4	11,500	0.0031			SRL
a-5	12,000	0.0032			SRL
a-6,7	10,000	0.0027			SRL
a-8,9	7,500	0.0020			SRL
a-10-13	2,500	0.0007			SRL
b	12,000	0.0204			EL
c-1,2	3,000	0.0043			SRL
d-1-3	5,000	0.0138			
e-1	4,000	0.0165			SRL
e-2	3,000	0.0124			SRL
e-3,4	1,200	0.0049			SRL
f	2,500	0.0097			
g	1,200	0.0038			
h	2,000	0.0072			
Q-a-1-4	16,000	0.0102§	0.0117‖		SC
a-5	15,000	0.0095§	0.0109‖		SC
a-6	10,000	0.0064§	0.0073‖		SC
b-1	13,333	0.0051	0.012		SRL
b-2,3	15,000	0.0057	0.013		SRL
b-4	16,000	0.0061	0.014		SRL
b-1	13,333	0.0020	0.0051		SRL
b-2,3	15,000	0.0022	0.0057		SRL
b-4	16,000	0.0024	0.0061		SRL
b-5,6	25,000	0.0037	0.0095‖	0.0053‖	SRL, SC & BR
c	1,200	0.0034	0.0163‖		SS
d	2,500	0.0051	0.0093‖	0.0136‖	SS
e-1,2	3,750	0.0036			SRL
f	1,200	0.0033§			
g	2,500	0.0032	0.0062‖	0.0106‖	FB
h-1	1,500	0.0019	0.0172‖		FB
h-2	1,500	0.0019	0.0172‖		FB
i	2,500	0.0032	0.0062	0.0106‖	FB
j	2,500	0.0025			SRL
R-a-1,2	4,000	0.0068		0.0187‖	
b	1,000	0.0064	0.029‖	0.0202‖	SRL
c	6,000	0.0101		0.0131‖	
d-1,2	15,000	0.0129	0.0188‖	0.0175‖	SC
d-3	6,250	0.0054	0.0078		SC

* Capital letter identifies the reporting utility. Lower case letter represents a particular location. Number identifies one furnace at that location.

† Points in these columns represent nonobjectionable flicker unless otherwise indicated.

‡ Nomenclature for this column:

SC = synchronous condenser
C = series capacitor
BR = buffer reactor
SRL = separate radial line
FB = flicker bus
ST = separate transformer
EL = extra transmission line
SS = separate substation
CS = capacitor in series with synchronous condenser

§ Borderline flicker.

‖ Objectionable flicker.

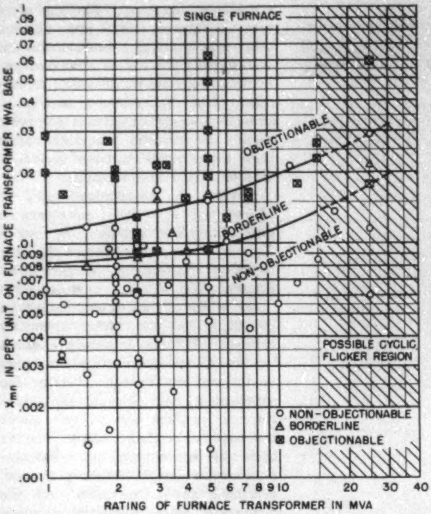
synchronous machines. For the usual case, the furnaces will be some distance away from generating stations, and the system reactance will be composed primarily of lines and transformers. Consequently the results shown in Figs. 3 through 5 would be affected only very slightly if transient reactance rather than subtransient reactance had been used for synchronous machines.

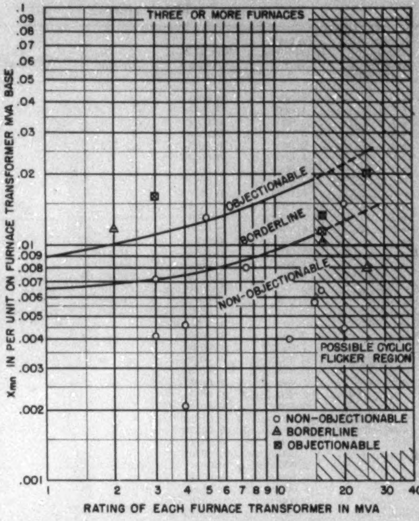
However, an installation might be planned for a location quite near a generating station bus, where the value of X_{mn} is appreciably affected by the choice of transient or subtransient reactance. The planning engineer must decide which of these reactances will be governing for the flicker frequency expected. For the extremely frequent rate, the effective reactance will probably be the transient reactance because the subtransient time constant of a synchronous machine will be quite short compared to the flicker period. For the cyclic rate either transient or subtransient reactance might be the effective one, depending upon the rapidity of the flicker rate. Once the effective reactance is selected, the curves in Figs. 3 through 5 can be used directly.

Early Attempts to Draw Limit Curves

The importance of having a large number of points upon which to base limit curves is well illustrated by one of the projects undertaken by the Working Group. Each of the operating companies was asked to attempt to draw limit curves based on only their own data. The limit curves submitted by the various companies are shown in Fig. 6. The difference between the highest and lowest curve in Fig. 6 is four to five times as wide as the band shown in Fig. 3. The reason for this great difference in bandwidth is easily understood, and in fact it formed one of the principal reasons for undertaking this industry survey.

The data of any one customer who has only a few arc furnace installations on his system will of necessity tend to yield a wide band between the limit curves. The fact that one company draws a very low curve below which satisfactory performance is assured does not necessarily mean that a value of X_{mn} above this curve will provide unsatisfactory performance. It may only reflect that the company has no furnaces operating satisfactorily above that line. Obtaining accurate limit curves requires a large number of installations in the marginal range. No one company will have a sufficient number of installations in this





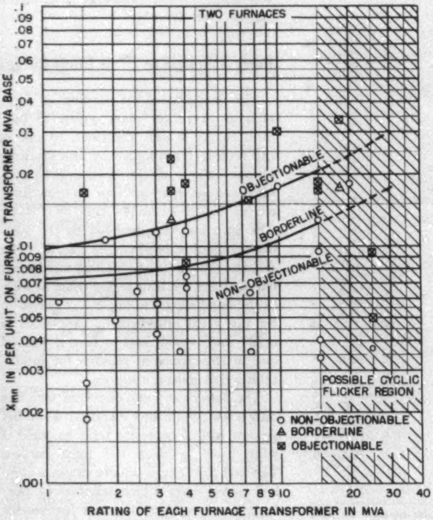


Fig. 3 (above left). Plot of X_{mn} from Table I as function of transformer rating for single-furnace installation. Limit curves for furnaces with electrode voltage of 350 volts or under

Fig. 4 (left). Plot of X_{mn} from Table I as function of transformer rating for two identical furnaces at same location. Limit curves for furnaces with electrode voltage of 350 volts or under

Fig. 5 (above right). Plot of X_{mn} from Table I as function of transformer rating for three or more identical furnaces at same location. Limit curves for furnaces with electrode voltage of 350 volts or under

range, and it is only when all the data compiled by the survey are considered that an accurate picture can be obtained of the true limit curves. The large amount of data in the survey does assure enough points in the marginal region to permit drawing accurate limit curves.

In the early stages of the survey an attempt was made to distinguish between X_{mn} limits for circuits supplying industrial areas and those supplying residential areas. This was subsequently abandoned in favor of the single, and perhaps somewhat wider, bands shown in Figs. 3, 4, and 5. However, the comments made by the reporting companies on this subject, given in the Appendix, were most illuminating.

Corrective Measures Employed

Column 8 of the survey sheet shown in Fig. 2 requested information about the corrective measures employed to control flicker at the various installations. This information is summarized in Table II. Correlation of corrective measures and size of the installation is shown in two ways. The first part of the table shows the corrective measures applied for various values of the largest furnace size at an installation; the latter part shows the corrective measures applied for various values of total furnace kva at an installation.

Table II shows that few of the smaller installations require special flicker corrective measures. As the installation

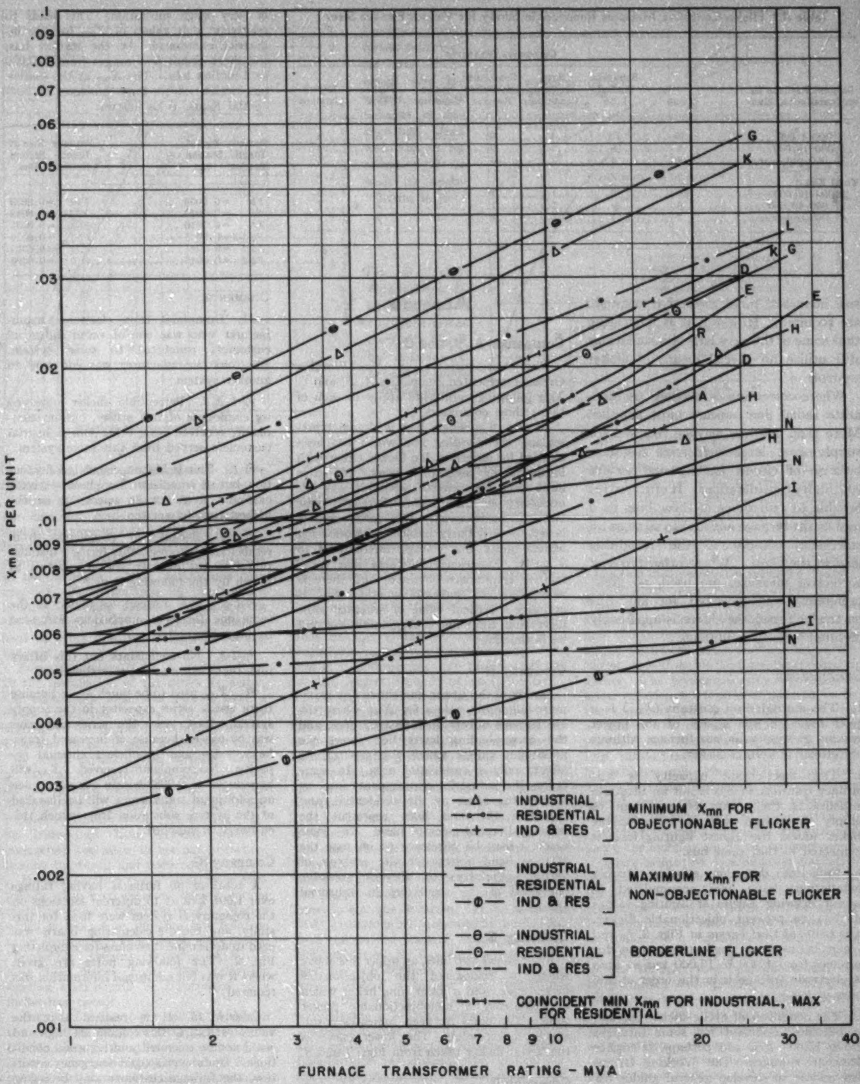


Fig. 6. Objectionable flicker limits as judged by ten reporting companies based on own data only

Table II. Flicker-Corrective Measures Reported in Survey for Various Furnace Sizes

Largest Furnace in an Installation, Kva	Corrective Measures						
	None	Separate Radial Line	Synchronous Condenser	Condenser and Buffer Reactor	Series Capacitor	Flicker Bus	Miscellaneous
1,000–4,000	39	13	1	0	0	3	2
5,000–15,000	9	8	1	2	1	2	1
16,000 and over	4	4	2	3	1	0	1
Total Kva							
1,000–10,000	39	12	1	1	1	3	2
11,000–40,000	6	9	1	1	0	2	1
41,000 and over	7	4	2	2	1	0	1

size increases, more corrective measures are required. However, it is interesting that some of the very largest installations still utilize no special means of flicker control.

Where correction is needed, the separate radial line appears most popular. More than one line may be required for supplying a large industrial customer because of circuit loading and service continuity considerations. It often proves feasible to utilize one of these lines as a feed to the furnace and also to serve as an emergency supply to the remaining load in the plant. Where other types of corrective measures are used, no great preference was indicated for any one particular type; the choice is apparently dictated by local conditions.

Conclusions

1. The mutual drop constant (X_{mn}) is a good index of the ability of the power system to supply an arc furnace without objectionable voltage flicker.

2. The short-circuit capacity of the furnace transformer bus is not an adequate measure of the ability of a system to supply a furnace without objectionable flicker unless the closest lighting load is connected to that same bus.

3. Sufficient data on existing furnace installations have been accumulated to permit drawing curves of limiting values of X_{mn} to prevent objectionable flicker. It is believed that curves in Figs. 3, 4, and 5 can be used as application curves for furnaces from 1,000 to 15,000 kva so long as electrode voltage is in the order of 350 volts or less.

4. The condition of cyclic flicker appears to become important for some furnaces above 15,000 kva and perhaps at higher electrode voltages. The Working Group was unable to develop general guides for these large furnaces, which should be studied as special cases.

5. The tolerable X_{mn} for 2-furnace installations is 10% to 15% under that for a single furnace.

6. Where special corrective measures are required, the separate radial line is used oftener than any other corrective means.

Appendix

Companies A, B, and C

The representative on the Working Group represented companies *A*, *B*, and *C*. The following comments apply to each of these three companies.

In Fig. 6 a single line is drawn representing the dividing line which has been applied for objectionable flicker either with industrial or residential users. No cases have been encountered where it was felt necessary to differentiate between these two classes of users. The curve used is based on previously published information which limits the voltage fluctuations to 2 volts. Experience indicates that satisfactory results are obtained, and there is very little to indicate whether this is actually a critical value or whether some increase in the value of X_{mn} might not still be permissible.

Company D

In Fig. 6 the upper line shows the maximum allowable values for X_{mn} where the arc furnace affects an industrial area, and the corresponding lower line shows the maximum values where the arc furnace affects only a residential area. In some cases, the industrial requirements may be almost the same as the residential, and therefore the two lines represent the industrial requirements band. In some cases it may be necessary to increase the voltage band setting of the controls of voltage regulators to prevent excessive operation of the regulators in industrial areas.

Company E

Fig. 6 shows two lines, an upper line above which is considered the objectionable flicker zone, and a lower line, below which is considered a nonobjectionable flicker zone. The band between indicates a tolerable flicker zone. The lower curve is the 2-volt flicker taken from Figs. 7 and 8.

Company F

The following notes refer to column 6 of Table I for company *F:*

The value of X_{mn} as given is the proper value for the electrically closest load. In these instances it represents other (from one to seven) industrial customers. In the system, these industrial customers are served on a 12-kv loop from a generating or very large substation. This leads to relatively high values of X_{mn} for the industrial customers. At the station bus, the short-circuit kva ranges from 500,000 to 2 million kva. The X_{mn} at the station bus, which affects large amounts of residential loads, is as follows:

Furnace Identification	X_{mn} at Station Bus	Furnace Identification	X_{mn} at Station Bus
Fb	=0.0056	*Fi-1*	=0.0023
Fe	=0.0010	*Fi-2*	=0.0012
Ff	=0.0010	*Fl-1, 2*	=0.0021
Fg-1-3	=0.0025	*Fl-3*	=0.0016
Fh-1	=0.0041	*Fl-4*	=0.0013
Fh-2	=0.0045	*F -0*	=0.0026

COMMENTS

Fb. Complaint from electronic manufacturer who was one of seven industrial customers connected to same system. Electronic manufacturer was cut over to another system.

Fg-1,2,3. Perceptible flicker observed by employees of this utility, but no complaints received from other two industrial customers served from this same system.

Fi-1. This is highest calculated fluctuation, but no complaints have been received; customer may not be operating on the highest transformer tap.

Fk-1-4. Because of complaints from residential customers, this furnace installation was cut over to another system as shown by the following item *F-1.*

Fl-1-4. This furnace customer is the same one formerly supplied as indicated by preceding note.

Fp-1-3. No complaints but this utility feels that this is at the tolerable limit.

Ft. X_{mn} used to be much lower because there was a series capacitor in the supply system. Last year, the series capacitor was by-passed because of increased transformer size and insufficient thermal capacity. No complaints received. X_{mn} will be reduced to 0.0083 within one year when an additional transformer will be installed at the system substation from which this customer is supplied.

Company G

A total of 36 furnaces having ratings over 1,001 kva at 15 different locations on the company *G* system were used for this study, and the d-c calculating board was used to determine the values for completing Fig. 2. The following notes are given where it was felt additional information was required:

Column 15. It is realized that the values of X_{mn} in this column are high and could not be tolerated under normal conditions. Under a prolonged emergency condition, the furnace customer may be asked to reduce the melt-down voltage in order to minimize flicker.

Furnace Location G-a-1 and 2. To prevent objectionable flicker at the time of the larger furnace installation, it was necessary to connect it to an existing 23-kv line in the area and operate the line radially from a bulk power supply station. The initial smaller furnace was supplied by the lines

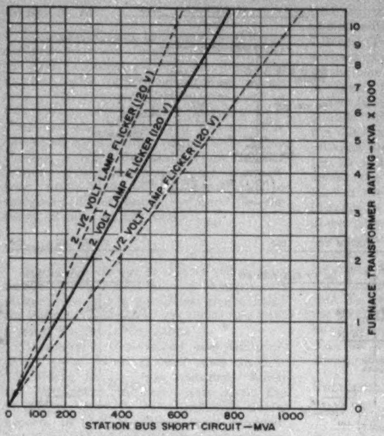
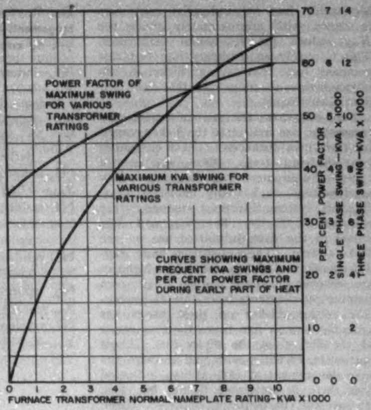


Fig. 7 (left). Furnace transformer rating versus bus short-circuit mva for various flicker voltages at furnace bus. Used with Fig. 8 by company E as basis for establishing limits

Fig. 8 (right). Approximate 3-phase arc furnace characteristics used with Fig. 7 by company E as basis for establishing limits



in the area, but was transferred to the radial line along with the larger furnace.

Furnace Location G-c-1 and 2. Before the customer could increase his furnace load to the present two 10,000-kva-capacity furnaces (rebuilt from 7,500 kva), a bulk power supply station was located adjacent to the customer. This station not only supplied system stiffness at the furnace location but also capacity for supplying the increased furnace load and for relieving load on other bulk power stations feeding the area.

Furnace Location G-d-1, 4, and 5. Prior to the installation of the series capacitor, the customer's load was served over a 69-kv radial line with a tap to a substation, which supplied emergency service to a small area through a 69/23-kv bank. The series capacitor was installed ahead of the tap to improve voltage regulation on the furnace customer's bus, but, incidentally, this improved the voltage flicker on the tap station so that the station could be transferred from normal stand-by emergency operation to normal operation, feeding industrial and residential loads.

Furnace Location G-g-2. Customer complaints have been received during normal operation from a television sending station which is located in the area but from no other customers. This column was checked "borderline" because the value of X_{mn} is near the flicker limit. The television sending station will be transferred to a flicker-free circuit.

Furnace Location G-j-1 and 2. Although the values of X_{mn} do not indicate flicker trouble, these furnaces have been at various times a source of complaints due to faulty operation or incorrect setting of the furnace limiter.

Furnace Location G-m. Coincident with the cut-in of this 5,000-kva furnace there were a number of complaints from residential customers in the area. Complaints ceased when an existing third line in the area, which had been operated open, was closed in to provide a stiffer system in the area.

On this system there are relatively few single-furnace installations, and these are small furnaces. As is indicated in the main body of this report, in plotting the data for 2-furnace installations where the furnaces are of different sizes, it is a question which size furnace to plot. If the furnace sizes are the same, there can be very little question. These remarks also apply to the 3-, 4- and 5-furnace installations.

Practically all of the furnace installations on this system started out with a single furnace, and it is this utility's opinion that it is more pertinent to plot the points for the largest furnace at each location. This would give the lower limit of the value of the short-circuit current required for the various sizes of furnaces. For example, if there is an 18,000-kva furnace satisfactorily working at one location, it is a foregone conclusion that furnaces of smaller size at the same location will operate satisfactorily.

It is believed that the major problem is to determine the maximum size of furnace that can be carried with the maximum X_{mn}. This having been determined, the result can be adjusted to show the reduction in the X_{mn} values for two, three, or more furnaces at one location. It is believed that a reasonable assumption could be that the second furnace would require a 20% reduction in X_{mn} as compared with that for one furnace; the third furnace would require a further 10% reduction in X_{mn} than for the first two, etc. Actually in analyzing a 3-phase kilowatt chart and a 3-phase kilovar chart of a single furnace, as compared with similar charts of a multiple-furnace installation, the charts for the multiple-furnace installation are smoothed out very materially as compared with the single-furnace charts. While load swings of a multiple-furnace installation were found to be higher than for a single furnace, the frequency of these higher load swings was much lower than for the lower load swings of a single furnace. It is questionable whether the additional furnaces contribute very much to the swing of the first furnace, particularly when it is considered that furnace load swings are usually single-phase and of short duration. While the swings of the two furnaces could theoretically occur simultaneously, actually in a majority of the overlaps the swing of one furnace would be on a different phase than that of the other furnace.

This company does not feel that it is in a position to show what is termed a "band" above which objectionable flicker would result and below which there would be no disturbing flicker. Therefore the lower line *G* has been drawn on Fig. 6, which might be described as the borderline between objectionable and nonobjectionable flicker. The position of this line is based upon experience which is considered to be the top limit (maximum that is safe) for certain of the furnace installations. For example, this utility is now serving a 4-kv unit substation practically from the same bus which supplies two 10,000-kva furnaces. There have been complaints from customers served from this 4-kv unit substation, therefore consideration is being given to providing a second source for the unit and using the existing source as an emergency.

There is another case wherein corrective equipment has been installed for normal use and it has been found that this corrective equipment could be taken out of service for maintenance and repairs without any complaints from lighting customers. Prior to taking the equipment out of service, it was thought that operation could not be permitted without it. Thus, this point establishes another upper limit. Another case where the data are shown plotted above the lower line is where, due to necessary system changes, the short-circuit current in the immediate vicinity of the furnace was reduced. A corrective plan was available but after operating without its use with no residential complaints, it has been postponed indefinitely. In this case there have been complaints from a television broadcasting station but not from other customers, and the station was transferred to a flicker-free supply in 1956.

The upper line *G* on Fig. 6 for the borderline between objectionable and nonob-

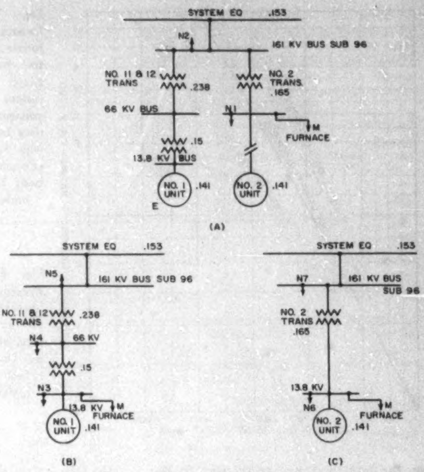


Fig. 9. System arrangements for analysis by company K

N1, industrial
$X_{mn}=0.0711$ objectionable
N2, residential
$X_{mn}=0.0289$ borderline
N3, industrial
$X_{mn}=0.028$ O.K.
N4, residential
$X_{mn}=0.020$ borderline
N5, residential
$X_{mn}=0.008$ O.K.
N6, industrial
$X_{mn}=0.0244$ O.K.
N7, residential
$X_{mn}=0.0117$ O.K.

jectionable line for industrial customers is shown with approximately twice the X_{mn} values for residential customers. This is purely an arbitrary line but this company feels that voltage flicker is less objectionable to industrial than to residential customers.

Furnace customers are industrial customers and are subject to the flicker caused by their own furnaces. These customers have drafting rooms, offices, and certain types of equipment requiring reasonably good voltage, as has any industrial user. On this system 2-furnace customers are supplied at 69 kv. The X_{mm} and X_{mn} values in the data submitted are on the 69-kv system. These customers have step-down banks to 11 kv to which their furnaces are connected. They supply their own lighting requirements from this 11-kv bus. The voltage flicker on these buses has probably four or five times the magnitude of the flicker on the 69-kv bus. These customers, recently contacted through the sales department of this utility, indicated that there is noticeable flicker, but that it has not particularly bothered them. One customer with a series capacitor in his 69-kv supply indicates that the flicker is not at all objectionable on his 11-kv bus when the capacitor is in service. However, with the capacitor out of service, it is objectionable but not too much so. This customer, identified as *G-d* in Table I, has never considered installing any corrective equipment on his 11-kv system to offset the furnace voltage flicker even though he does have offices and drafting rooms supplied therefrom.

The other 69-kv customer identified as *G-b* on Table I indicates that while the flicker is noticeable, he has never considered doing anything to correct it.

Company H

Fig. 6 shows two bands formed by an upper, a middle, and a lower line. The upper band defines the limit for industrial installations, and the lower for the residential user. These bands are based on arc furnace installations having a 40% reactance.

The bands indicate the flicker conditions as they exist. Installation *H-a* is in good condition operating as a 40% reactance installation; but when the secondary voltage is raised, which, in effect, reduces the percentage reactance of the installation, flicker becomes noticeable. Installation *H-b* causes flicker which is tolerated. Installation *H-c* has noticeable flicker but does not appear to be objectionable in a residential area. The three boundary lines from upper line to middle to lower represent respectively approximately $1^1/_2$%, 2%, and $2^1/_2$% flicker conditions based on Figs. 7 and 8, wherein X_{mn} is obtained by dividing "furnace transformer rating" by "station bus short circuit" and plotting it with its appropriate "furnace transformer rating."

Observe the lower limit of the residential band as drawn in Fig. 6, which shows at the 10,000-kva transformer size X_{mn} to be 0.01 and thus the system short-circuit kva to be approximately one million. Assuming that the furnace transformer and system accounts for 0.10 per-unit X, and the leads, brackets, and electrodes account for 0.30 per-unit X on a 10,000-kva base with the furnace operating on the 275-volt taps, a 3-phase swing from 85% power factor to 24% power factor gives an approximate 1.5% flicker voltage at the mutual bus. When this same furnace is operated on the 350-volt taps, the furnace transformer and system account for 0.10 per-unit X; but the leads, brackets, and electrodes now account for only 0.185 per-unit X on the 10,000-kva base, and a 3-phase swing from 85% to a 24% power factor gives an approximate 2.0% flicker voltage at the mutual bus. To maintain the flicker voltage at 1.5% while operating on the 350-volt taps, the system short-circuit kva would have to be increased to approximately 1,333,000, and this in turn reduces the X_{mn} to 0.0075. It appears from this example that the location of the lower limit line as plotted through 0.01 X_{mn} for the 10,000-kva transformer is appropriate for a 0.40 per-unit reactance installation; but for the 0.285 per-unit reactance condition, this lower limit would have to be 0.0075 X_{mn} (at 10,000 kva) if it is to be equally applicable to this reduced reactance condition.

It appears that the reactance of the load, and the reactance in the feed line to it from the bus at which X_{mn} is calculated, is an important factor in this problem and must be given consideration in the final results. There is some question as to whether or not the high values of X_{mn}, column 6, normal system operation, might show up as having a high lead and furnace transformer reactance or that companies that operate normally with those nigh values of X_{mn} are in the fortunate position of being able to interpret the question of objectionable flicker much more liberally than other companies who feel it necessary to make those expenditures which assure a lower X_{mn} for the same furnace size.

The feeling of this company is that it is necessary to base the determination of suitable system conditions for arc furnace installations on other factors besides the kva rating of furnace transformer. The two associated factors which must be considered are, first, the effective reactance of the furnace load which depends on secondary voltage and, second, the character of other loads served from the common bus with respect to their susceptibility to voltage variations.

Company I

The two lines in Fig. 6 represent this company's opinion of the maximum and minimum limits of X_{mn} required to determine objectionable and unobjectionable flicker voltages from arc furnace operation. This utility company's system serves several large industrial customers whose electrical needs are such that voltage flickers are extremely undesirable, and for this reason it is believed that a single set of limitations should apply to industrial and residential customers alike.

Each of the two furnaces served by this system are served at 66 kv and each has synchronous condenser-buffer reactor corrective equipment which was sized by the customer to provide for operation of future additional furnaces. As of March 1955, the additional furnaces were not installed.

Company K

See Fig. 9.

Company L

Company *L* has no cases of objectionable flicker and all values of X_{mn} reported in column 6 represent satisfactory operating conditions. This being the case, it is possible only to draw a line as indicated in Fig. 6 below which all values of X_{mn} represent nonobjectionable flicker conditions. Due to the sparseness of data, the true slope of the curve is rather problematical.

Company M

Company *M* has had very limited experience with arc furnace loads, and, as the data presented denote, the experience has been with very small kilowatt ratings only. The two arc furnaces connected to this system are located electrically very close to major generation busses. The effects of these loads are negligible as to causing voltage fluctuations on lighting busses. For this reason, this company feels unqualified by its experience to designate flicker boundary lines for the range of arc furnace loads.

Company N

Applicable curves in Fig. 6 show four lines to indicate the limits of objectionable flicker; the upper two lines give a band for industrial users and the lower two lines a band for residential users.

Company *N* experienced one real complaint about flicker in 1949 when the first 15,000-kva furnace was installed. When the 20,000-kva furnace, and later the 25,000-kva furnace were installed, a few reports of voltage flicker were received, but these did not persist after the furnaces had been operated for a few months.

During the years in which the furnace load was increasing, the system was also increasing, so that the short-circuit capacity at the substation bus from which this customer is supplied is now greater than necessary to dispose of the flicker problem.

Company O

Company *O* feels that they cannot establish bandwidth or limits, since there is only one furnace in operation on the system. However, it has been determined that for this single furnace in operation, objectionable flicker occurs with X_{mn} of 0.0222. This value of X_{mn} is obtained without corrective measures being applied to the transmission line which supplies the 5,000-kva furnace.

The furnace is installed at the end of a 40-mile 69-kv line. Compensation for the one transmission line serving this load provides proper voltage levels and stability for an intermediate load, but leaves unsolved the problem of serving additional load at the furnace end of the line. After installation of a higher-capacity bus at the power plant and the location of additional intermediate load, a change in operating procedure was required. It was found by removing compensation and making a second line parallel with the original one, that X_{mn} was reduced low enough to keep the flicker within acceptable limits. It can be inferred, however, that the limit of X_{mn} for this particular furnace is somewhere between a value of 0.0065 and 0.022.

Company P

No comments.

Company Q

Furnaces *Q-b-5* and *-6* are supplied by radial lines from a power plant so that X_{mn} is composed of generator plus transformer reactances. The transient reactance was used in planning the corrective equipment, whereas the subtransient reactance is required in this survey. The value of X_{mn} for *Q-b-5* or *-6* would be about 50% higher if the transient value had been used.

Company R

The applicable curve of Fig. 6 outlines the approximate system short-circuit capacity (expressed as X_{mn} on the furnace transformer rated kva base) required to avoid objectionable flicker for a single furnace. The curves for two, three, or more arc furnace installations, at the same location, can be translated from the characteristics reported by Ramsauer and Treweek.[1] Our data are not sufficiently extensive to establish a bandwidth to these curves beyond which the furnace operation can be definitely assured as satisfactory or unsatisfactory. Also, there are no circuits that are used exclusively to supply either industrial or residential customers, and therefore one limitation applies to both load classifications.

Arc furnace *R-b* in Table I is supplied from a 1,000-kva furnace transformer connected to a radial 12-kv line almost three miles long, and from which no other load is supplied except the steel plant lighting. Flicker is of objectionable magnitude at the plant, but no complaint has been made. Residential customers are supplied from the substation facilities which are also the source of supply to the line to this industrial customer.

The three furnaces listed as *R-d-1*, *-2*, and *-3* are presently supplied from the 25-cycle system from which no critical lighting customers are supplied. However, between the latter part of 1941 until the last quarter of 1945, these furnaces were operated from a 60-cycle system, and this is the period covered by the data in the tabulation.

Reference

1. **System Capacity Required for 3-Phase Arc Furnaces**, **Oliver Ramsaur, J. E. Treweek.** *AIEE Transactions*, vol. 73, pt. III-B, Dec. 1954, pp. 1373–83.

Discussion

C. Concordia and **L. G. Levoy** (General Electric Company, Schenectady, N. Y.): This fine report is a useful and informative guide. There are, however, a few points which we believe need further discussion. We appreciate the practical reasons for plotting the data against name-plate kva, but it must be remembered that furnaces with the same name-plate rating may have different reactances and different maximum tap voltages, and consequently will perform differently.

Regarding so-called cyclic flicker on large furnaces with higher secondary voltages, we wish to call attention to some recent experimental work which indicates generally greater arc stability and fewer surges in higher voltage furnaces.[1]

In some cases, cyclic variations of current have been observed in small furnaces near the beginning of the melt but, since the supply system is usually large compared with the furnace size, this often goes unnoticed. True cyclic variations (variations recurring at a constant frequency) might be caused by improper adjustment of the electrode regulators, by vibration of the electrodes and supporting structure, by ripples on the surface of the melt, or similar disturbances of a periodic nature. Most other fast and slow variations seem to be random or aperiodic in nature, and may be caused by the changing geometry and composition of the arc. Such changes result from movements of the hot spots on the electrodes or on the surface of the melt or by the changing geometry of the melt.

Few of the smaller installations require special flicker corrective measures, and as the size increases more corrective measures may be required. On the other hand, since some of the very largest installations require no corrective means, it is evident that the important thing is not the size of the furnace but its relative size with respect to the power supply.

The expanding use of arc furnaces will be accelerated by providing adequate power supplies which can serve the furnaces without special corrective means. For economic reasons, large arc furnaces will be used. The shaded area shown on Figs. 3, 4, and 5 over 15,000 kva may tend unnecessarily to discourage the use of these large furnaces. Since some larger furnaces operate without the cyclic flicker it is important to determine what is wrong with those that do have it and to correct it.

We would appreciate it if the committee would clarify the reported cases of trouble of this kind by providing the following additional information:

1. Is the reported cyclic flicker truly cyclic or is it fast, random flicker?
2. Does this disturbance persist throughout the melt-down period, or does it occur for only a few minutes at the beginning of the melt?
3. What is the effect of starting on a reduced voltage tap for the first 5 minutes of operation? Some operators find that by reducing voltage for the first few minutes until the electrodes and work are heated, flicker is greatly reduced during this interval. Then they go to the top tap for rapid melt-down.
4. Is the fast flicker persisting in these furnaces at present, or was this observed when the furnaces were first placed in operation? We have observed that as the operators become more skillful in charging and operating the furnaces, and when the controls are properly adjusted, many of the flicker problems which show up initially are minimized or eliminated.
5. It is our understanding that Figs. 2, 3, 4, and 5 are intended as application guides based upon actual service experience. On this basis it would not seem proper to

include the dashed "possible cyclic flicker" region since we understand that this cyclic flicker is based not upon actual complaint experience but simply upon an extrapolation of flicker sensitivity curves. Is this understanding correct? Satisfactory operation of a large furnace with an extremely small X_{mn} would not in itself violate the borderline curves as originally drawn, and it therefore seems somewhat inconsistent and confusing to bias the upper region of the curves.

It appears that large electric arc furnaces will be very important to the steel industry in the future. It is therefore essential to learn how to serve them economically without excessive disturbance to the system. We believe that this report provides a step forward toward that objective.

REFERENCE

1. EXPERIMENTAL RESULTS WITH HOLLOW ELECTRODES IN ELECTRIC STEEL FURNACES, W. E. Schwabe. *Iron and Steel Engineer*, Pittsburgh, Pa., June 1957, pp. 84–92.

W. E. Schwabe (National Carbon Company, Niagara Falls, N. Y.): Investigations of this kind are very valuable in view of the future development of electric furnace practice and its further expansion within the steel producing industry.

Several years ago, the National Carbon Company started an extensive research project concerning furnace arcs in the laboratory and in the trade. One of the major objectives was to study the physical conditions that caused the electric arc to be unsteady during the melting operation in the furnace and, if possible, to improve arcing conditions.

Our company being one of the results of this work, we can state today that the phenomenon of erraticity of the arc can, in general terms, be explained. Arc unsteadiness is caused by interaction of magnetic forces acting upon the arc and forces originating in the arc. In combination with random changes of the geometry of the melt surface and variation of the ionization in the arc zone, variations of the arc current are mainly identical with variations of the instantaneous arc length. Therefore the problem of eliminating or reducing current swings reduces to the necessity of keeping the arc in a more steady position.

As to the definition of a common criterion for lighting flicker, I should like to suggest that the name-plate kva rating of the transformer be replaced by the kva of the furnace load average under operational conditions during the melt-down. The name-plate rating of the transformer is often not identical with the furnace load, especially with older installations. Furthermore, the terminology of "extreme frequent flicker" and "cyclic flicker" may be misleading: "extreme frequent" occurrence of a phenomenon is rather relative, and the so-called cyclic flicker is not of a true cyclic nature. Perhaps better and more precise terms can be found.

We have investigated arcs in steel furnaces from a laboratory size of 100 kw to a large furnace of 130-ton capacity and 25,000 to 27,000 kw. Over this entire range we find that essentially two types of current fluctuations were present:

Type A: Load swings that occurred less frequently than one per second with a relatively great intensity ("extreme frequent" swings).

Type B: Quasi-cyclic variations of a somewhat smaller magnitude, but occurring at a rate of 3 to 7 per second (cyclic swings).

The types are intermixed, and they occur especially during the first 15 to 20 minutes of a melt-down with cold scrap. Details were reported at the 1956 Association of Iron and Steel Engineers convention in Cleveland.

We designed and built an electronic counting device which, installed in the arc circuit, counted the number of load swings of both types, and we followed in detail several hundred heats at small, medium, and large furnaces with this instrument. The result is that both types *A* and *B* swings occur less frequently in a large furnace of 25,000 kw than in a medium size furnace of 19,000 kw or a small furnace of 5,000 kw. This is not in agreement with the extrapolated curve of the report.

Type *B* (or cyclic) current swings are the result of arc flare in connection with turbulence of the melt under the electrode. Large furnaces operate with higher electrode voltages, and consequently the gap between electrode and melt is larger. The influence of wave ripple, as a quasi-cyclic modulator of the arc length, decreases with increased gap or arc voltage. Therefore the problem of lighting flicker in relative terms is less critical with large furnaces.

We made extensive recordings of the line voltage variations caused by furnace operation. For our own information, standards of visible and objectionable flicker were established. Ninety per cent of the persons involved in this statistical-psychological test had no objection to voltage dips of 0.75% at a rate of five per second, at 60- to 75-watt light bulbs. Details on this subject will be reported in the near future.

It should be stressed that the problem of lighting flicker is very closely related to the problem of furnace efficiency. Deviations of the arc current from a desired magnitude, such as the furnace optimum current, reduce the input of power to the furnace. Arc steadiness, reduction of lighting flicker, and improved furnace operation go hand in hand.

As a practical means to improve these conditions, we developed and manufactured special electrodes, for example, hollow electrodes for experimental work. Hollow electrodes reduce the load swings considerably. Our measurements indicate an average reduction in the order of 50%. This is due to a focusing of the arc under the tip of hollow electrodes and a reduction of side flashing of the arc. Furthermore, we investigated electrodes with cores of various compositions, which create a hollow tip and which, however, did not show the same improvement.

More work must be done and co-ordination of all efforts is essential because the trend in the electric steel business is to build and use even larger furnaces.

C. E. Quick (The Detroit Edison Company, Detroit, Mich.): This survey presents some very interesting information, which can be of considerable assistance to engineers. There are, however, two points on which caution should be observed in the use of the survey report. The data, as explained in the second paragraph of the report, make no provision for variations in the allowable limits for voltage flicker between the several power companies. Further, as explained in the third paragraph, due to the difficulty in predicting the single-phase swings of a furnace, it is assumed that furnaces of a given size will all have the same magnitude and rate of swing. For furnaces of 15,000-kva and under, the normal electrode voltage is 350 volts or less. The single-phase swings at an extremely frequent rate will normally be the critical factor for these furnaces. The limit for this rate of swing is quite widely accepted as 1.5% or 2 volts (on a 120-volt base). Further, since the range of electrode voltage for these furnaces is relatively small, the magnitudes of the swings are undoubtedly comparable for the several furnaces of any given size. The curves shown in Figs. 4 and 5 for transformer sizes under 15,000 kva are probably quite safe to use.

For furnaces larger than 15,000 kva the first increases in size were accomplished simply by increasing the current, through the use of larger electrodes, lower impedance lead circuits and transformers, etc. Recently there has been an increasing trend to higher electrode voltages. Our experience has shown that with an increase in electrode voltage there is an increase in the magnitude of single-phase swings at a cyclic rate of four to six times per second.

On 25,000-kva furnaces with 510 electrode volts melting loose scrap, the magnitude of the cyclic swings is approximately 60% of the extremely frequent swings. Since the normal allowable flicker voltage limit, at a cyclic rate is 0.5% or less (one third of the flicker limit for extremely frequent swings), the cyclic flicker may cause trouble on a system where the extremely frequent flicker is allowable.

In view of this, it is strongly recommended that furnaces above 20,000 mva (megavolt-amperes) be installed only where the X_{mn} is 0.005 or less, unless complete calculations are made concerning the voltage flicker which will be experienced at both the extremely frequent rate and at the cyclic rate.

B. L. Lloyd: We wish to express the appreciation of the Working Group to the discussers for their valuable and informative discussions.

Mr. Concordia and Mr. Levoy point out that name-plate kva alone does not completely describe the swings to which a furnace is subjected. This, of course, is true and if a particular installation differs markedly from the typical installation in reactance, lead length, etc., some correction factor should be applied to the limits of X_{mn} indicated in this paper.

The remainder of this discussion asks a number of specific questions concerning experience with "cyclic" swings. These questions were transmitted to company *Q* and the following replies received:

1. The cyclic flicker, measured with oscillograms by one company on two

25-mva furnaces operating at 510 electrode volts and melting loose scrap, was shown to be true cyclic flicker, ranging from 4 to 12 cycles' duration occurring four to six times per second.

2. The cyclic flicker occurs during the first 30 minutes of melt-down.

3. Starting on a reduced voltage tap will reduce the cyclic flicker. However, the reduced voltage tap must be continued until a pool of molten steel has formed. When the melt is loose scrap, the period of reduced voltage must be at least 30 minutes to avoid cyclic flicker.

4. Checks made with a Brush oscillograph more than a year after the 25-mva furnaces began operations showed the cyclic flicker to be present at essentially the same magnitude and rate.

5. Oscillograms show that 16-mva furnaces have single-phase swings at both a cyclic rate and at an extremely frequent rate with such magnitudes that the extremely frequent swings are the controlling factor for design of system arrangements. Similar oscillograms for the 25-mva furnaces with 510 electrode volts show a relative increase in magnitude of the swings at a cyclic rate with the result that the cyclic swings become the controlling factor for system design. As previously stated in the survey report, it is probable that the high electrode voltage has much to do with the increase in relative magnitude of swings at a cyclic rate. The total magnitude of these swings remains a function of both the electrode voltage and the size of the furnace.

Mr. Schwabe's discussion provides considerably interesting and valuable background information on why furnaces experience the load swings that they do. Concerning the terminology of "extremely frequent" and "cyclic" flicker, while more accurately descriptive terms could be devised, these terms are well accepted and understood by the industry and probably will continue to be used.

Mr. Schwabe states that his tests, indicating that less frequent swings occur with the large furnaces than with small furnaces, are in disagreement with the extrapolated curves in the report. We should first point out that these curves indicate a level of annoyance, and result from not only a frequency of occurrence but also from the magnitude of the swings. Further, the extrapolated curves would indicate smaller relative swings for the large furnaces than for the small, since higher values of X_{mn} would be permitted for the large furnaces. The extrapolated curves would appear to be in agreement with Mr. Schwabe's tests. However, the report of higher cyclic swings on the large furnaces with higher electrode voltage casts uncertainty on the extrapolated curves.

The discussion by Mr. Quick and the other discussers express individual viewpoints on the importance of cyclic swings for the larger furnaces. As was pointed out in the paper, the purpose of this report was to present results of a survey of experience of companies which have arc furnaces in operation on their systems. Adhering to this purpose, the Working Group can only state that the survey did not encompass a sufficient number of these very large furnaces to permit establishing limits in this range. The various viewpoints of the discussers are most welcome.

The Stability of Aircraft D-C Power Systems with Inverter Loads

R. E. THOMAS
NONMEMBER AIEE

THE ANALYSIS presented in this paper is intended to provide a basis for investigating the effects of inverter loads on the stability of aircraft d-c power systems. There have been occasions when energizing a particular inverter has produced an unstable system. In the past, such a situation has been resolved, with varying degrees of success, by an arbitrary change in the time response of one or more of the offending elements. The problem has been subjected to several intuitive explanations, probably the most popular of which is that the time constant of the inverter and the time constant of the generator were the same. While such reasoning is intuitively correct, it is scientifically meaningless, and provides little or no guidance to the system designer. Since inverter-induced instability is a system problem, it must be analyzed on a system basis. The objective of this paper is to perform a system analysis, and to point out the measurable properties of the system which may be used to predict stability.

Nomenclature

$E_d(t)$ = voltage drop across generator internal impedance
$E_m(t)$ = inverter motor back emf (electromotive force)
$E_0(t)$ = generator internal generated voltage
$E_t(t)$ = generator terminal voltage
f_p = frequency of maximum generator-regulator closed-loop gain
$G_1(p)$ = generator-regulator open-loop transfer function
$G_1'(p)$ = generator-regulator closed-loop transfer function
$G_2(p)$ = armature circuit transfer function
$G_g(p)$ = generator open-loop transfer function
$G_r(p)$ = regulator open-loop transfer function
$G_T(p)$ = system open-loop transfer function
$G_T'(p)$ = system closed-loop transfer function
$I_f(t)$ = inverter motor field current
$I_m(t)$ = inverter motor armature current
$I_0(t)$ = generator armature current
I_m = inverter rotary section polar moment of inertia
K_c = inverter speed control gain
K_g = generator gain
K_r = regulator gain
K_s = generator compounding constant
K_m = inverter motor torque constant
K_v = inverter motor voltage constant
L_f = inverter motor field inductance
L_e = generator armature inductance
M_p = maximum generator-regulator closed-loop gain
$N(t)$ = inverter speed
p = differential operator
$R_f(t)$ = inverter motor field resistance
R_m = inverter motor armature resistance
R_0 = generator internal resistance
Y_m' = normalized inverter input admittance
Z_L = load impedance
Z_m = inverter input impedance
Z_e = generator internal impedance
β_n = undamped natural frequency of inverter input impedance
β_n' = undamped natural frequency of armature circuit transfer function with inverter load
δ = damping ratio of generator-regulator closed-loop transfer function
δ' = damping ratio of system closed-loop transfer function with resistive load
τ_r = regulator time constant
τ_g = generator time constant
σ = damping ratio of inverter input impedance
σ' = damping ratio of armature circuit transfer function with inverter load
ω_n = undamped natural frequency of generator-regulator closed-loop transfer function
ω_n' = undamped natural frequency of system closed-loop transfer function with resistive load

Formulation of a Mathematical Model

One of the first steps in system analysis is the formulation of a mathematical model. In such a formulating every effort should be made to include only

---

Paper 57-478, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted January 15, 1957; made available for printing March 15, 1957.

R. E. THOMAS, lieutenant, is at Wright Air Development Center, Dayton, Ohio.

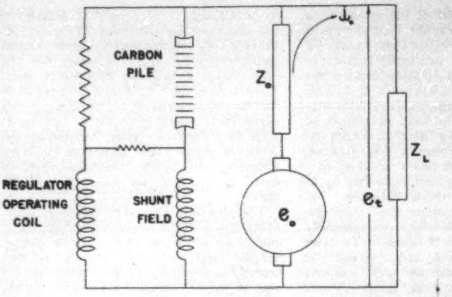


Fig. 1. Schematic diagram of system

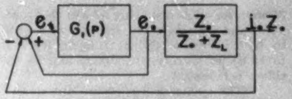


Fig. 2. Block diagram of system with inner loop open

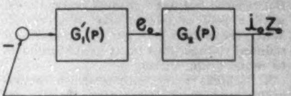


Fig. 3. Block diagram of system with inner loop closed

those conditions and properties which are essential to the immediate problem. In the study of nonlinear systems it is often advisable to begin at least with a linearized model of the system, and to evaluate the validity of linearization, by comparing results obtained from the linear model with those obtained from the actual physical situation.

The aircraft d-c system is not unique among physical systems in that it contains nonlinearities. A comprehensive analysis covering all of the possible conditions and properties which might affect d-c system stability must of necessity include the effects of nonlinearities. Such an analysis leads one into a realm of intangibles, and becomes extremely complex clear-cut answers to positive questions. However, by restricting the application of the mathematical model to the one specific problem of d-c system stability with inverter loads, it is possible to obtain satisfactory answers from the study of a linearized system.

The aircraft d-c system may be linearized by restricting the analysis to small disturbances about a fixed operating point. It is customary to speak of analysis based upon such a linearization as "the method of small motions."[1] The small-motion approach is often used in the analysis of nonlinear physical systems. For example, the so-called linearized equations of motion, used by the aerodynamicist in studying the dynamic longitudinal and lateral stability of airplanes, are based upon small disturbances about a fixed operating point. In fact, previous studies[2,3] have shown that the dynamic performance of aircraft d-c generators and carbon-pile regulators may be predicted. When studies are restricted to small disturbances. The use of small-motion linearization to predict the stability of nonlinear systems are investigated in reference 4. It suffices to say here that the small-motion approach will adequately predict stability, provided the nonlinearities of the system are continuous, and the disturbances from the fixed operating point are small. Any attempt to extend the analysis to areas where these conditions are not satisfied should be undertaken with caution.

In analysing the d-c system by this technique, there is need for initial clarity in the definition of the variables to be studied. The variables are defined as the small disturbances, while the operating points become parameters.* The system to be analyzed consists of a shunt-wound d-c generator, a carbon-pile voltage regulator, and a load; it is shown schematically in Fig. 1. The application of simple circuit analysis to the circuit of Fig. 1 yields the following basic equations:

$$\frac{e_0(t)}{e_t(t)}=G_1(p) \tag{1}$$

$$\frac{i_0(t)Z_0}{e_0(t)}=\frac{Z_0}{Z_0-K_s+Z_L} \tag{2}$$

$$e_t(t)=e_0(t)-i_0(t)Z_0 \tag{3}$$

In deriving these equations, and in the development that follows, the following standard assumptions are made:

1. The input impedance to the field circuit is high in comparison with the impedance of the armature circuit.
2. The generator speed is constant.
3. Commutation is perfect.
4. At the frequencies of interest, the effects of eddy currents may be neglected.

In addition, it may be assumed that the effects of compounding may be neglected, in which case the term K_s in equation 2 may be neglected. These conditions are assumed in the development that follows, although the effects of compounding are discussed later in the paper.

Since the system is to be analyzed by means of frequency response, and by the application of feedback control theory, it is more convenient to represent the system in block-diagram form, as shown in Fig. 2. For convenience of measurement, it is necessary to close the inner feedback loop, since to measure the open-loop transfer function of conventional carbon-pile regulators is difficult. The result of closing this loop is shown in Fig. 3. The two transfer functions in Fig. 3 are defined as

$$G_1'(p)=\frac{G_1(p)}{1-G_1(p)} \tag{4}$$

$$G_2(p)=\frac{Z_0}{Z_0+Z_L} \tag{5}$$

Where $G_1(p)$ is the generator-regulator closed-loop transfer function at no-load, and $G_2(p)$ is termed the armature circuit transfer function.

Thus the system has been reduced to two transfer functions in series with direct negative feedback. However, in the actual physical situation the output cannot be measured directly, nor can the feedback loop be broken to determine the open-loop transfer function. It is thus necessary to study the transfer functions separately, and to compute the open-loop transfer function. For this three properties of the system must be known: the closed-loop transfer function of the generator-regulator combination, the input impedance of the load, and the internal impedance of the generator.

In reference 2 it is shown that the generator and the regulator may each be represented by a gain and a time constant, as follows:

$$G_r(p)=\frac{K_r}{\tau_R p+1} \tag{6}$$

$$G_g(p)=-\frac{K_g}{\tau_g p+1} \tag{7}$$

Hence the closed-loop generator-regulator transfer function is given by equation

---

* Throughout this paper lower-case letters with the argument t (time) represent the disturbance, while upper-case letters with the argument t represent the total variable, i.e., operating point plus disturbance.

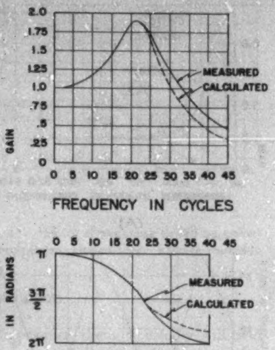


Fig. 4. Typical generator-regulator closed-loop transfer function

8 as follows

$$G_1'(p) = -\frac{K_rK_g/\tau_r\tau_g}{p^2+\dfrac{\tau_r+\tau_g}{\tau_r\tau_g}p+\dfrac{1+K_rK_g}{\tau_r\tau_g}} \qquad (8)$$

Since the open-loop gain K_rK_g is much greater than 1, it is convenient to define the closed-loop transfer function in terms of its undamped, natural frequency and damping ratio, thus:

$$G_1'(p) = -\frac{\omega_n^2}{p^2+2\delta\omega_n p+\omega_n^2} \qquad (9)$$

$$\omega_n = \sqrt{\frac{1+K_rK_g}{\tau_r\tau_g}} \qquad (10)$$

$$\delta = \sqrt{\frac{(\tau_r+\tau_g)^2}{4\tau_r\tau_g(1+K_rK_g)}} \qquad (11)$$

Equation 9 will be recognized as describing a classical, linear, second-order system. The gain of such a system is 1 at low frequencies, rises to a resonant peak the height of which is determined solely by the damping ratio, and then falls off asymptotically to 0 at high frequencies. The maximum gain is often referred to as peak magnification, and is given the symbol M_p. The relationship between peak magnification and damping ratio is given by equation 12.

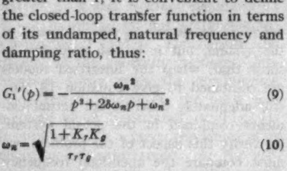


Fig. 5. Typical normalized generator internal admittance

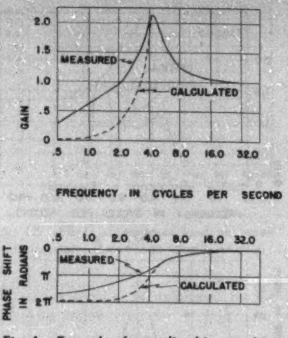


Fig. 6. Example of normalized inverter input admittance

$$M_p = \frac{1}{2\delta\sqrt{1-\delta^2}} \qquad (12)$$

The relationship between the frequency at which the peak gain occurs and the undamped natural frequency is given by equation 13.

$$\omega_n = \frac{\omega_f}{\sqrt{1-2\delta^2}} \qquad (13)$$

By knowing M_p and the frequency at which it occurs, the generator-regulator closed-loop transfer function may be defined in terms of damping ratio and undamped natural frequency. High peak gains are indicative of lightly damped transient modes, and under these conditions the frequency of the peak will be within a few cycles of the natural transient frequency. Much can be learned about the general character of a d-c system by examining the generator-regulator closed-loop transfer function.

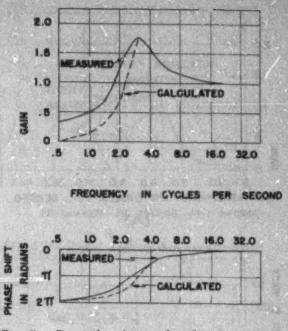


Fig. 7. Example of normalized inverter input admittance

To determine the armature circuit transfer function $G_2(p)$, it is necessary to know the input impedance of the inverter and the internal impedance of the generator. The derivation of the linearized equation for the input impedance of an inverter is given in the Appendix. There it is shown that, in the frequency range of interest, the input impedance of the inverter may be approximated by

$$Z_m = R_m\left(\frac{p^2+2\sigma\beta_n p+\beta_n^2}{p^2}\right) \qquad (14)$$

For generators with commutating fields and interpole windings, it is usually possible to neglect the inductive portion of the generator internal impedance. Hence the armature circuit transfer function for inverter load is

$$G_2(p) = \frac{R_0}{R_0+R_m}\left(\frac{p^2}{p^2+2\sigma'\beta_n'p+\beta_n'^2}\right) \qquad (15)$$

where

$$\beta_n' = \beta_n\sqrt{\frac{R_m}{R_0+R_m}} \qquad (16)$$

$$\sigma' = \sigma\sqrt{\frac{R_m}{R_0+R_m}} \qquad (17)$$

Thus by combining equations 9 and 16 a linearized mathematical model of the system open-loop transfer function is obtained as given by equation 18.

$$G_T(p) = \frac{R_0}{R_0+R_m}\left(\frac{-\omega_n^2p^2}{(p^2+2\delta\omega_n p+\omega_n^2)(p^2+2\sigma'\beta_n'p+\beta_n'^2)}\right) \qquad (18)$$

Study and Verification of the Model

Equation 18 is, of course, an approximation, since a number of idealizations were made in its derivation. However, it must be remembered that equation 18 is to be used to obtain some criteria for the stability of the system, and that therefore one should be primarily concerned with the degree to which it fits the actual physical situation, in the frequency range in which the stability of the system is determined. To determine this frequency range, one must fall back on one of the properties of the actual system, namely, that the generator-regulator natutal frequency is usually several octaves above the inverter natural frequency. This condition immediately suggests a system whose gain is characterized by two resonant peaks with a relatively flat valley between the two, and with low gain at high and low frequency.

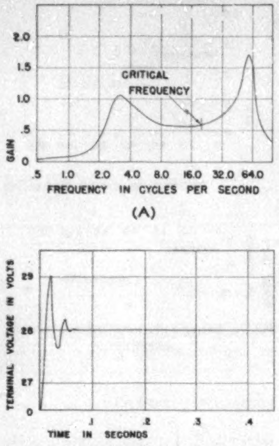


Fig. 8. Example of system response

A—Open-loop frequency response
B—Closed-loop transient response

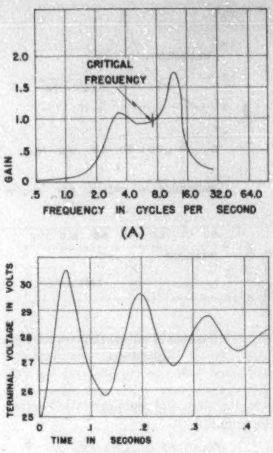


Fig. 9. Example of system response

A—Open-loop frequency response
B—Closed-loop transient response

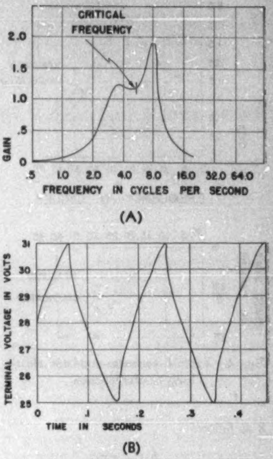


Fig. 10. Example of system response

A—Open-loop frequency response
B—Closed-loop transient response

Further, the phase shift would be zero at low frequency, pass through 90 degrees in the neighborhood of the first peak, and reach 180 degrees somewhere in the valley between the two peaks. Since the gain in the region of 180 degrees phase shift is the criterion for judging the stability of a system, it is apparent that equation 18 must be a good approximation for frequencies greater than the inverter natural frequency, but less than the generator-regulator natural frequency. That is, in order to be useful, equation 18 must adequately describe the actual physical system in the following frequency range:

$$\beta_n' \leq \omega \leq \omega_n \tag{19}$$

This frequency range will be termed the system mid-frequency range.

Furthermore, if there is considerable difference between the inverter natural frequency and the generator-regulator natural frequency, the mid-frequency gain will approach

$$|G_T| \rightarrow \frac{R_0}{R_0+R_m} \tag{20}$$

and since this ratio must always be considerably less than 1, the system will be stable. If, however, the natural frequencies approach each other, it is possible for the mid-frequency gain to increase and exceed 1, resulting in an unstable system. Thus it is to be expected that one of the fundamental requirements for the stability of d-c system with inverter loads will be the separation of the two natural frequencies.

The extent to which the individual elements of the actual system are described by the mathematical model can be seen by examining Figs. 4, 5, 6, and 7. Fig 4 is a typical example of the measured generator-regulator closed-loop transfer function compared to that calculated from equation 9. For frequencies less than the peak frequencies, that is, in the mid-frequency range, this element is adequately described by equation 9. Fig. 5 is an example of the measurement of generator internal impedance. Since the normalized generator internal admittance indicates a half-power frequency in the neighborhood of 80 cps (cycles per second), and since the system critical frequency will usually lie below 20 cps, it is safe to assume that the inductive component of generator internal impedance may be neglected. Figs. 6 and 7 are examples of measured inverter input admittance compared to that calculated from equation 14. The significance of these figures is discussed in more detail in the Appendix; however, it can be seen that in the mid-frequency range, that is, for frequencies greater than the peak frequency, equation 14 is adequate.

It has been demonstrated that each of the linearized models is individually adequate in describing the elements of the system, but it is also necessary to show that, when the linearized models are combined to give equation 18 they will adequately describe the actual elements combined in the actual system. To verify this aspect of the problem, we must compare the open-loop frequency response of the system obtained by measuring each of the elements with the closed-loop transient response, because the actual open-loop frequency response cannot be directly measured. Figs. 8 through 10 are the results of tests run on d-c systems with inverter loads. The approach was by testing systems with widely different generator-regulator natural frequencies, and observing the closed-loop system transient response. The tests show that as the gain of the open-loop system at the critical frequency increases, there is a deterioration of system stability. Significantly, the increased gain at the critical frequency results from moving the two resonant peaks of the open-loop transfer function closer together. It is therefore concluded that the problem of d-c system stability with inverter loads may be studied with the use of a linearized mathematical model; and that the fundamental requirement for the stability of such a system is that the generator-regulator natural frequency and the inverter natural frequency must be separated.

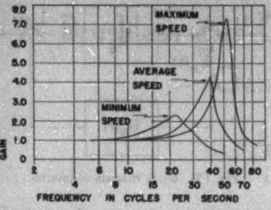


Fig. 11. Effect of generator speed on generator-regulator closed-loop transfer function

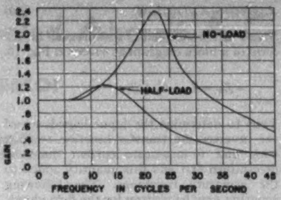


Fig. 12. Effect of resistance load on generator-regulator closed-loop transfer function

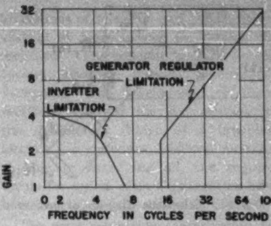


Fig. 13. Signal performance requirements

Discussion

In practice, there is usually sufficient separation of the inverter natural frequency and the generator-regulator natural frequency to insure stability. This occurs by accident due to the nature of the system. The inverter rotary section moment of inertia and motor field inductance are relatively large energy-storing elements, and hence limit inverter natural frequencies to the range of 2 to 6 cps. In addition, generator-regulator combinations designed to meet current space and accuracy requirements do not, in general, exhibit lightly damped low-frequency modes. The natural frequency of the generator-regulator combination varies with the combination being tested. However, the most pronounced effect (illustrated in Fig. 11) is obtained with a given combination by changing generator speed. As the generator speed is increased, both the peak gain and natural frequency increase greatly. This means that at high generator speeds the d-c system tends to have lightly damped high-frequency natural modes, while at low speeds the natural modes are more damped and lower in frequency. This effect is primarily due to the fact that the generator gain increases with speed; from an examination of equations 10 and 11 it can be seen that increasing generator gain increases the undamped natural frequency and decreases the damping ratio. From the standpoint of system stability with inverter loads, the lightly damped high-frequency modes can be tolerated, and since the low-frequency modes usually have a reasonable amount of damping, a vast majority of generator, regulator, inverter combinations will be compatible. However, there is nothing in current specifications to insure this condition.

The basic requirement for the stability of the system is the separation of the natural frequencies. The amount of separation required is influenced by the following factors:

1. The minimum mid-frequency gain given by equation 20. This gain may be regarded as a measure of the "stiffness" of the bus. The stiffer the bus, the lower the minimum mid-frequency gain, and the less the separation of natural frequencies is required. The power transfer and voltage drop requirements are usually sufficient to maintain the minimum mid-frequency gain within reasonable limits.

2. The effect of resistive load. For resistive load the armature circuit transfer function becomes

$$G_2=\frac{R_0}{R_0+R_L} \tag{21}$$

In this case the closed-loop system transfer function is given by

$$G_T'=-\frac{\omega_n'^2}{p^2+2\delta'\omega_n'p+\omega_n'^2} \tag{22}$$

where

$$\delta'=\sqrt{\frac{(\tau_r+\tau_g)^2}{4\tau_r\tau_g\left[K_rK_g\left(1-\frac{R_0}{R_0+R_L}\right)+1\right]}} \tag{23}$$

$$\omega_n'=\sqrt{\frac{K_rK_g\left(1-\frac{R_0}{R_0+R_L}\right)+1}{\tau_r\tau_g}} \tag{24}$$

An example of the effect of resistance load upon a typical system is shown in Fig. 12. From an examination of this figure and equations 10, 11, 23, and 24, it can be seen that effect of resistance load is to increase the damping ratio, and to decrease the natural frequency. The former effect is, of course, stabilizing, while the latter is destabilizing from the standpoint of system stability with inverter load. The increase in damping is usually the predominant effect, and the over-all result of resistance load is usually stabilizing.

3. Generator compounding has a pronounced effect on system stability with inverter load. If the generator compounding is not negligible, then the mid-frequency gain, the armature circuit natural frequency, and the damping ratio are given by the following equations:

$$|G_T|\rightarrow\frac{R_0}{R_0-K_s+R_m} \tag{25}$$

$$\beta_n'=\beta_n\sqrt{\frac{R_m}{R_0-K_s+R_m}} \tag{26}$$

$$\sigma'=\sigma\sqrt{\frac{R_m}{R_0-K_s+R_m}} \tag{27}$$

If the generator is overcompounded, K_s is positive and the result is an increase in the mid-frequency gain, the natural frequency, and the damping ratio. All of these effects are detrimental to system stability with inverter loads. However, current generator specifications preclude overcompounding, except perhaps at high generator speeds, in which case there is usually sufficient separation of the natural frequencies due to the effect of generator speed upon the generator-regulator natural frequency.

Summary

The principal factor which affects the stability of d-c systems with inverter loads is the separation of natural frequencies. To this must be added several secondary effects which may be divided into two broad classes: those which result from the power transfer properties of the system, and those which are related to the signal properties of the system. The power transfer properties are covered, or at least adequately restricted, by existing performance requirements. The signal performance requirements, those which insure the separation of natural frequencies, are not specified. Therefore, the signal performance requirements of Fig. 13 are suggested. These were obtained by correlating the results of an analog computer study with measurements obtained from the laboratory. The requirements of Fig. 13 may be listed as follows:

1. The maximum generator-regulator closed-loop gain shall be limited in magnitude according to the frequency at which it occurs by the curve indicated in Fig. 13.

2. The generator-regulator must maintain a gain of 1 to at least 14 cps if there is no peak in the frequency response.

3. The maximum normalized input admittance of the inverter shall be limited in magnitude according to the frequency at

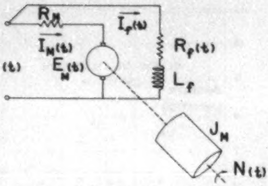


Fig. 14. Schematic diagram of inverter

which it occurs by the curve indicated in Fig. 13.

4. The normalized input admittance must have reached 1 by 14 cps if it does not exhibit any peak.

It must be understood that these requirements are the antithesis of mathematical logic in that they are the sufficient conditions but are not the necessary conditions. That is, the fact that a particular generator-regulator combination does not meet the requirements of Fig. 13 does not necessarily mean that it will be incompatible with inverters which do meet the requirements. It should also be pointed out that if a system designer specifies the range of frequencies in which certain equipment may operate, then steps must be taken to insure that the system does not create any periodic disturbances within this frequency range, for example, those which might arise from cyclic loads or mechanical vibration.

The purpose of the foregoing analysis has been to define the problem of d-c system stability with inverter loads. A mathematical model was derived in terms of measurable properties of the system; by correlating results obtained from the mathematical model with laboratory tests, it was possible to verify the model and to obtain the suggested requirements of Fig. 13. The importance of the work lies in the background which it provides for extension to other systems in which devices exhibiting lightly damped, oscillatory natural modes are interconnected merely for the purpose of transferring power.

Appendix. Inverter Input Impedance

The linear expression for the input impedance may be obtained from the inverter shown schematically in Fig. 14. The following equations describe the behavior of the inverter:

$$\frac{E(t)-E_m(t)}{R_m}=I_m(t) \tag{28}$$

$$E_m(t)=K_vI_f(t)N(t) \tag{29}$$

$$J_mpN(t)=K_mI_f(t)I_m(t) \tag{30}$$

These equations may be expressed in the form of a block diagram as shown in Fig. 15.

$$\frac{E(t)}{I_m(t)}=Z_m=R_m+G_m(p) \tag{31}$$

Thus to determine the input impedance of the inverter the transfer function relating back emf to armature current must be determined first and this quantity added to the armature resistance. The following equation may be written for the field circuit

$$E(t)=I_f(t)R_f(t)+L_fpI_f(t) \tag{32}$$

or by definition

$$E+e(t)=[I_f+i_f(t)][R_f+r_f(t)]+L_fpi_f(t) \tag{33}$$

If the products of the disturbances are neglected, then

$$e(t)=I_fr_f(t)+(R_f+L_fp)i_f(t) \tag{34}$$

If it is assumed that the time delays in the inverter speed control are very short compared to the motor field and the mechanical time constants, then

$$v_f(t)=K_cn(t) \tag{35}$$

If it is further assumed that for small disturbances the variations in applied voltage have little effect upon field current compared to that resulting from the variation in carbon pile resistance, then equation 36 relates changes in field current to speed.

$$-K_cI_fn(t)=(R_f+L_fp)i_f(t) \tag{36}$$

By definition equation 29 may be written as

$$E_m+e_m(t)=K_v[I_f+i_f(t)][N+n(t)] \tag{37}$$

The change in back emf is given by

$$e_m(t)=K_vI_fn(t)+K_vNi_f(t) \tag{38}$$

Similarly the change in torque may be obtained from equation 30 as

$$J_mpn(t)=K_mI_mi_f(t)+K_mI_fi_m(t) \tag{39}$$

Equations 36, 38, and 39 are three linear equations which describe the transfer function G_m for small disturbances about a fixed operating point. Equation 36 relates motor field current and speed. Equation 38 relates back emf to field current and speed. Equation 39 relates torque to field current and armature current. Solving these equations for $G_m(p)$ yields

$$G_m(p)=\frac{K_m(\tau_mp+1)}{p^2+a_mp+b_m} \tag{40}$$

where

$$K_m=\frac{K_cK_TI_f(R_f-K_cN)}{J_mL_f} \tag{41}$$

$$\tau_m=\frac{L_f}{R_f-K_cN} \tag{42}$$

$$a_m=R_f/L_f \tag{43}$$

$$b_m=\frac{K_cK_mI_mI_f}{J_mL_f} \tag{44}$$

Combining equations 31 and 40 yields an expression for the inverter input impedance equation 45

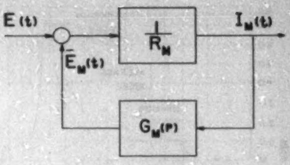


Fig. 15. Block diagram of inverter

$$Z_m=R_m\left[\frac{p^2+\left(a_m+\frac{K_m\tau_m}{R_m}\right)p+(b_m+1/R_m)}{p^2+a_mp+b_m}\right] \tag{45}$$

Thus, the input impedance is given by a quotient of two second-order polynomials. Actual measurements of inverter input impedance indicate that, in general, the roots of the denominator of equations 45 are very small compared to the roots of the numerator, and that, within the frequency range of interest, the input impedance may be approximated by

$$Z_m=R_m\left[\frac{p^2+2\sigma\beta_np+\beta_n^2}{p^2}\right] \tag{46}$$

where

$$\beta_n=\sqrt{\frac{b_mR_m+1}{R_m}} \tag{47}$$

$$\sigma=\sqrt{\frac{(a_mR_m+K_m\tau_m)^2}{4R_m(b_mR_m+1)}} \tag{48}$$

Figs. 6 and 7 are measurements of inverter input impedance. For convenience they are plotted as normalized input admittance which is defined as

$$Y_m'=\frac{R_m}{Z_m}=\frac{p^2}{p^2+2\sigma\beta_np+\beta_n^2} \tag{49}$$

These figures also show the frequency range in which equation 46 approximates the actual measurements. At low frequencies the neglected terms of the denominator of equation 45 are still important. However, in the mid-frequency range equation 46 is a good approximation, and this range is important when considering the effect of inverters on the stability of the system.

References

1. INTRODUCTION TO NONLINEAR MECHANICS (book), **N. Minorsky.** J. W. Edwards, Ann Arbor, Mich., 1947.

2. TRANSIENT ANALYSIS OF A VOLTAGE-REGULATED AIRCRAFT D-C SYSTEM, **Donald G. Scorgie.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1318–23.

3. DYNAMIC CHARACTERISTICS OF CARBON PILE VOLTAGE REGULATORS, **Robert L. Mills.** *Ibid.*, pp. 1347–53.

4. PROBLÈME GÉNÉRAL DE LA STABILITÉ DU MOUVEMENT (book), **M. A. Liapounoff.** Annals of Mathematical Studies, vol. 17, Princeton University Press, Princeton, N. J.

A Brushless Air-Cooled Aircraft A-C Generator

R. E. SMITH
ASSOCIATE MEMBER AIEE

Synopsis: The quantity and complexity of electric apparatus on aircraft have increased so much that maintenance of this apparatus has become a serious problem. On future aircraft, which will have even more complex requirements for electric apparatus, programs must continue to be initiated to provide more reliability and simplicity of this apparatus if the aircraft are to be kept in flying readiness. Generators, the source of electric power for the aircraft, are among the most important components of the electric system. Some of the construction details and the means by which a brushless air-cooled aircraft a-c generator can contribute to the reliability and simplicity of an electric system are analyzed in this paper.

Discussion

ADVANTAGES OF A BRUSHLESS GENERATOR

Several advantages of the brushless generator are created directly by the elimination of brushes, and may be listed as follows:

1. There are no brush wear problems; therefore the altitude and environmental capabilities of the generator are increased.
2. Brush run-in has been eliminated and results in a reduction of required overhaul facilities and time.
3. Without brush wear, there is no carbon dust to collect on the windings and form low-resistance creepage paths.
4. Slip rings, commutator, and brush-holder rigging have been replaced by a single rectifier assembly, resulting in a shorter machine. Since bracket windows for inspecting and replacing brushes are not required, the bracket can be made in the form of a solid tube. Both of these factors make for a generator which is much stiffer mechanically and therefore better suited to air frame vibrational ambients.
5. There are no variations in brush neutral settings to affect exciter compounding or its resultant effect on voltage regulation and stability.
6. There will be no voltage oscillations if water exists in the generator cooling air. On conventional, brush-type generators, the water forms an occasional film between the brushes and the commutator, or slip rings. This film causes variations in current flow through the brushes which can result in violent generator output voltage oscillations.
7. Radio noise filters are unnecessary because there is no brush sparking to create radio noise.
8. Brush life testing is eliminated, which may be an advantage to the customer as well as to the generator manufacturer, since there is then much additional time for the use of altitude chambers to study cooling problems.

The use of an a-c exciter in a brushless generator makes possible further advantages:

1. Voltage regulator design is simplified because the range of excitation required by the a-c exciter over the generator temperature range can be greatly reduced from that of a conventional generator with d-c exciter.
2. Field flashing can be eliminated because sufficient exciter residual voltage for main generator voltage build-up can be readily provided by placing permanent magnet poles in the a-c exciter stator.

GENERATOR DESIGN

Class *C* requirements for generators require that they operate at higher altitudes and temperatures than ever before.[1] The higher temperatures and altitudes impose many problems in obtaining satisfactory brush life,[2–5] which, together with the advantages, caused the decision to build brushless generators. Some of the more important design features of a brushless 40-kva a-c generator are discussed in this paper.

Obviously, the most important problem in designing a brushless generator was that of obtaining rectifiers which were capable of delivering the necessary power under the space, weight, and cooling-air requirements of class *C* aircraft a-c generators. The lack of suitable rectifiers had been the only factor in the past which prevented the satisfactory design of production brushless generators. In 1951 a brushless 60-kva generator was built with selenium rectifiers, but was unsatisfactory for aircraft application because of the power limitations of selenium and the excessive amount of space needed for cooling the rectifiers.[6] The generator was very satisfactory, however, from a performance viewpoint. The development of silicon rectifiers removed the last obstacle to the satisfactory design of a brushless aircraft a-c generator. Silicon rectifiers capable of satisfying all the requirements for use in class *C* generators have been in production for the past 2 years.

The most promising configuration for a brushless generator had been determined as one which was similar to a conventional a-c generator, but which had the commutator, slip rings, and brushes replaced by a rotating rectifier assembly, and the d-c exciter replaced by an a-c exciter. It was further found that the feeding of a 3-phase exciter to a 3-phase full-wave rectifier assembly gave the greatest power output for the least generator size and weight. This determination was due to experience on an oil-cooled brushless 40-kva aircraft generator described in a previously presented paper.[7] That paper also gave other means which had been investigated for designing a brushless generator. Fig. 1 shows the schematic diagram of a brushless rotating-rectifier generator. Figs. 2, 3, and 4 show the rotating-rectifier assembly, generator cutaway, and complete generator.

With the basic circuitry established, the next major problem was determining where the rectifier assembly should be located in the generator. The first location considered was on the shaft outside diameter, but mounting it there would have added too much length and overhung moment to the generator. The next location considered, and the one selected, was inside the shaft. Since this space is normally wasted on conventional generators, and is dimensionally ideal for mounting the rectifiers, cooling calculations were made, and indicated that, with a simple cooling-fin construction rectifier, heating would be no problem under class *C* cooling conditions. These conditions were, blast cooling air temperatures up to 120 C (degrees centigrade), and a maximum allowed total head pressure drop across the generator of 11 inches of water. The calculations have been justified by testing. That the rectifiers are inside the shaft, has an additional advantage of subjecting them to a minimal centrifugal force.

Elimination of the requirement for field flashing was easily accomplished in the a-c exciter. Small, permanent-magnet poles were placed between the main poles, and being so located, the

---

Paper 57-468, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 11, 1957; made available for printing March 15, 1957.

R. E. SMITH is with the Westinghouse Electric Corporation, Lima, Ohio.

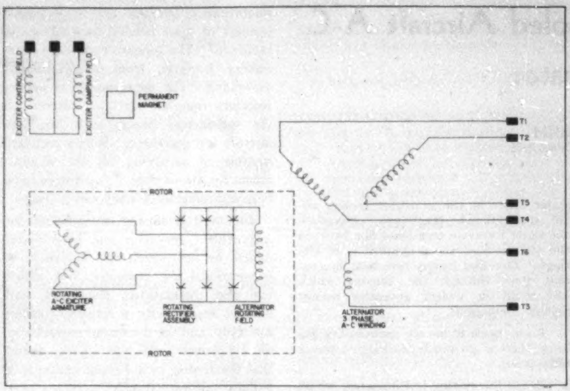

EXCITER CONTROL FIELD
EXCITER DAMPING FIELD
PERMANENT MAGNET
T1
T2
T5
T4
T6
T3
ROTOR
ROTATING A-C EXCITER ARMATURE
ROTATING RECTIFIER ASSEMBLY
ALTERNATOR ROTATING FIELD
ALTERNATOR 3 PHASE A-C WINDING
ROTOR

Fig. 1. Schematic diagram of brushless a-c generator

added space and weight to accommodate them was negligible. On conventional d-c exciters this space is occupied by interpoles, and thus prevents the existence of any convenient means for mounting the permanent magnets.

The permanent-magnet poles in the exciter stator generate voltage in the exciter armature as the generator is brought up to speed. The armature voltage produces main field current, resulting in a small generator output voltage. This output voltage is just sufficient to start current flowing from the regulator to the exciter field. This field current increases the exciter armature voltage, and the build-up process continues until rated generator voltage is reached.

The permanent-magnet poles are installed in a manner to cause the exciter armature reaction mmf (magnetomotive force) (resulting from the initial flow of exciter armature current) to polarize the main poles in a direction proper for aiding the regulator in starting the build-up process. When the generator is up to rated voltage, the exciter armature reaction mmf is in opposition to the permanent magnet mmf, and therefore almost completely cancels the effect of the permanent magnets.

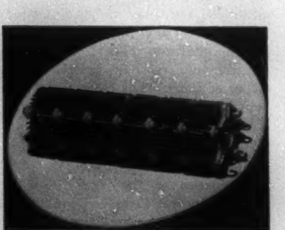


Fig. 2. Rotating rectifier assembly

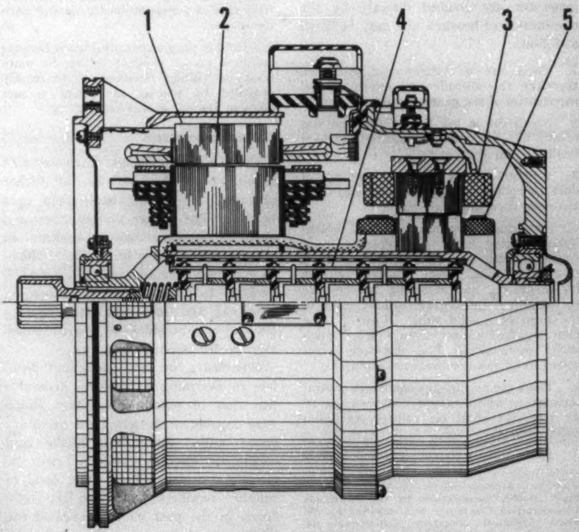


Fig. 3. Cutaway of brushless a-c generator

1. 3-phase a-c stator
2. Main rotating field
3. Exciter stator
4. Rotating rectifier assembly
5. 3-phase a-c exciter armature

Rectifier Characteristics

The silicon rectifiers used in the subject 40-kva brushless generator have half-inch junction diameters, and are rated at 20 amperes d-c average with no blast-cooling air at room temperature. This ampere rating means little, however, under blast-cooling conditions. The only limitation on the current rating is that the junction temperature remain below 200 C. Each of the rectifiers must pass a 1-minute air-cooled power test of 170 amperes d-c average. This would be equivalent to a rectifier bridge assembly output of 510 amperes d-c average. When installed in the generator, the maximum current which the assembly must provide is 120 amperes d-c average. This occurs under conditions of a 3-phase short circuit, during which the generator must deliver a minimum of 300% rated current for 15 seconds. Fig. 5 shows the cooling characteristics of the rectifier assembly when installed in the generator. It can be noted from the small temperature rises that the rectifiers, as installed in the generator, can readily be cooled under any class *C* operating condition, since the maximum air-in temperature for

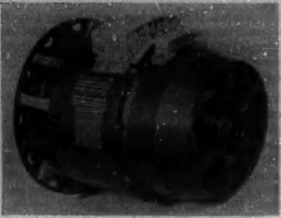


Fig. 4. Complete assembly of brushless a-c generator

class *C* conditions is 120 C and at this temperature 11 inches total head of cooling air is allowed. The reverse voltage rating of the rectifiers is 200 volts peak. The maximum reverse voltage which they encounter in the generator is 65 volts peak, and occurs upon application of a 3-phase short circuit at the generator terminals. This is a transient peak, and the voltage decays to 45 volts maximum after this peak.

Each rectifier has 40 milliamperes maximum peak reverse leakage, with 150 volts peak applied in an ambient of 180 C and each has a maximum forward drop of 1 volt with 50 amperes peak at room temperature. Both the leakage and forward drop are of main concern in the rectifier heating which they create.

Exciter Characteristics

The 3-phase a-c exciter was designed to have a high synchronous reactance, that is, its excitation requirements are more dependent upon its load current than on its load voltage. It can be considered a current amplifier, rather than a voltage amplifier such as a conventional d-c exciter. The purpose of this design characteristic was to simplify the matching of the generator to the regulator over the class *C* temperature range. Over this range the resistance of the alternator field, which is the exciter load, can change by a factor of 3 to 1. In a conventional exciter, the excitation required is almost completely dependent upon the load voltage, since armature reaction is mostly cancelled by compensating windings. Therefore, for a given alternator field current, the range of excitation required by a conventional exciter must vary by a factor of almost 3 to 1. The a-c exciter, on the other hand, can be readily designed to have the armature reaction effectively swamp out the effects of alternator field resistance variations. Fig. 6 shows the comparative ranges of excitation required by a brushless generator a-c exciter and by a conventional generator d-c exciter, for rated load over the class *C* temperature range. The narrow excitation range of the a-c exciter greatly simplifies the problem of matching the generator to the regulator, by reducing the required current range of the regulator magnetic amplifier.

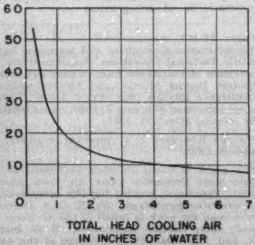


Fig. 5. Rectifier cooling characteristics at rated load

The excitation requirements of the a-c exciter are also much more consistent from generator to generator, and will remain more consistent over the operating life than they will for a conventional generator. The reason for the improved consistency is that there are no brush neutral setting, or brush contact film resistance, variations.

The damping field in the exciter stator is used to improve stability and prevent voltage regulator overcorrections for changes in generator output voltage. It sends a signal to the regulator which is proportional to the rate of change of flux in the exciter and is in opposition to the regulator current output change which caused the change in flux. On conventional generators, this same function is usually accomplished by the regulator sensing the rate of change of exciter output voltage at the commutator brushes.

Conclusions

The data presented in this paper have led to several conclusions. Where not obvious, the conclusions have been verified by testing.

1. A brushless generator is more nearly environment-free than a conventional generator.

2. A brushless generator will have more consistent performance than a conventional generator.

3. Maintenance is simplified because of the reduced number of parts compared to a conventional generator.

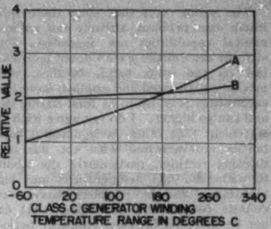


Fig. 6. Excitation requirements over class C temperature range

A—Conventional generator
B—Brushless generator

4. Reliability has been increased by elimination of items such as brushes, commutators, and slip rings which can have inconsistent performance, and by incorporating simple provisions for insured voltage build-up.

5. A brushless generator is more easily matched to the voltage regulator because the range of excitation required by its a-c exciter over the temperature range is much less than that required by the d-c exciter of a conventional generator.

References

1. Aircraft Generators. *Military Specification, MIL-G-6099A*, Aeronautical Standards Group, Mar. 25, 1957.

2. High-Altitude Brush Problem, D. Ramadanoff, S. W. Glass. *AIEE Transactions (Electrical Engineering)*, vol. 63, Nov. 1944, pp. 825–30.

3. Treatment of High Altitude Brushes by Application of Metallic Halides, Howard M. Elsey. *Ibid.*, vol. 64, Aug. 1945, pp. 576–79.

4. Air Humidity and Brush Contact Drop: The Effect of Water Vapor, Sulfur Dioxide, and Hydrogen Sulfide, Howard M. Elsey, Lawrence E. Moberly, John L. Johnson. *Ibid.*, vol. 73, pt. III-B, Dec. 1954, pp. 1383–89.

5. Effects of Silicone Vapor on Brush Wear, J. Marsden, Robert H. Savage. *Ibid.*, vol. 67, pt. II, 1948, pp. 1186–90.

6. An Alternator Without Brushes. *Westinghouse Engineer*, Pittsburgh, Pa., vol. 13, Mar. 1953, pp. 78–79.

7. An Oil-Cooled A-C Generator for High-Speed High-Altitude Aircraft, H. J. Braun, W. J. Shilling. *AIEE Transactions*, vol. 74, pt. II, 1955 (Jan. 1956 section), pp. 456–60.

◆

Discussion

M. H. Lowe (General Electric Company, Erie, Pa.): The author has presented a very interesting approach to the problem of design of an a-c generator for use in present-day aircraft. We must certainly all agree that this a-c generator is an improvement over the conventional a-c generator.

There are some points made in the paper with which I would like to take issue, and several questions I would like to ask.

The author says that since there is no brush wear problem, altitude and environmental capabilities are increased. Pure altitude capabilities are no doubt increased, but on the other hand, brushes, when properly engineered and applied, have been successfully used up to at least 80,000 feet, and can go higher. I do not agree with the statement that other environmental capabilities are increased with the use of a rotating rectifier, particularly the ability to withstand high temperatures and high vibrations. Silicon rectifiers have very definite temperature limitations that can not be exceeded for even transient periods without failing the rectifier. This weakness is not shared by brushes. Brushes are able to withstand high-temperature transients, such as occur in many aircraft installations where zero air-flow or reverse air-flow conditions exist during normal operations. In this respect I believe that the reliability of a rotating rectifier machine is much less than a statically excited machine using slip ring brushes only.

The author gives as an advantage the elimination of brush life testing. I feel the requirement for brush life testing is simply being replaced by a more extensive requirement for rectifier reliability testing. Brush life is simple to test, there is a great deal of test history to draw upon, and brush life can be fairly accurately predicted by extrapolating from shorter tests. On the other hand, rectifiers can fail in many ways, there is much less test history showing their reliability, and short-life tests can not be extrapolated to give predicted life. The life of a rectifier does not gradually decrease as does the length of a brush; one cannot look at a rectifier to determine whether it has 200 hours or 5 minutes remaining to its life. Thus preventive maintenance becomes guesswork.

I have two questions. Is the machine capable of operating under a ground idle condition of no air flow or more specifically can the machine meet the self-cooled requirement of *MIL-G-6099A* (see reference 1 of the paper) requiring 100% load at 40 C with no cooling air flow? Second, are the permanent magnets used to obtain build-up adequate to give control power under short-circuit conditions or does this power have to be obtained in another way?

R. E. Smith: I agree with Mr. Lowe in that brushes can possibly be developed to be used at very high altitudes. However, based upon past experience, the development of brushes has lagged behind the new altitude requirements being imposed upon generators. Brushes have been shown to be sensitive to various oil vapors and other gaseous fumes which can be present in the generator cooling air. The hermetically sealed rectifiers have no altitude restrictions and are not affected by these vapors and other contaminants.

Temperature limitations are no problem for even the existing production rectifiers. If the rectifiers are applied so that under any operation condition they do not exceed a given temperature, then there is no temperature restriction on the generator. It has been determined that it is far easier to cool the rectifiers in a high-temperature generator than it is to cool the windings. The rectifiers are cooled just as well with reverse air flow as they are with flow in the normal direction. Under conditions of no blast air, some type of fan must be used in any type of generator to cool the windings. It has been shown that the fan need be no larger to cool the rectifiers also. I would like to summarize two types of the many tests that have been run on brushless generators to show their reliability.

1. With the generator running at rated load and with a cooling air temperature of 120 C, which is the maximum required for class *C* generators, the generator was run for 400 hours with 11 inches H_2O cooling air pressure head, which is the head allowed for class *C* generators. The head was then reduced to 6 inches H_2O and the generator run for 100 hours. Finally the head was reduced to 3 inches H_2O and the generator run for 10 hours. At the completion of these tests there was no change in the rectifier characteristics and there were no failures.
2. A total of over 300 short circuits was applied to three ratings of brushless generators without a change in the rectifier characteristics and without a single failure.

The permanent magnets in the exciter stator are used for voltage build-up only. If it is desired to obtain control power from the generator then a small permanent magnet generator could be added on the shaft similar to that used on the static exciter generators. This would give a generator of approximately the same length as a conventional generator.

In summary, I must take issue with the statement that brushes will consistently give good performance at high altitudes and that their performance can be accurately predicted by extrapolating from short-time tests. Because of brush-wear problems and other advantages of a brushless generator given in the paper, I must conclude that a brushless generator is more reliable than a generator with brushes.

Prediction of Transient and Equilibrium Component Temperatures and Evaluation of Cumulative Insulation Damage of Blast-Cooled Machines

WALTER ROBINSON
NONMEMBER AIEE

IN THIS PAPER several aspects of correlating thermal test data of blast-cooled aircraft generators are reviewed. A graphical presentation of machine characteristics is introduced, which may be used in predicting steady-state and transient machine temperatures. Methods are discussed for determination of machine heat capacity and for calculation of transient component temperatures. Supporting experimental data are presented. Relationships for the calculation of cumulative insulation damage are developed and their applications suggested.

Nomenclature

a, A, b, C = constants
B = variable temperature coefficient of heat generation rate, Btu/hr–F (per hour–degrees Fahrenheit)
D = cumulative damage, dimensionless
G = heat generation rate constant, Btu/hr–F
H = heat transfer factor, Btu/hr–F
I = load current, amperes
K = heat capacity, Btu/F
L = insulation life, hr
N = speed of rotation, rpm
p = exhaust air pressure, inches Hg (mercury)
Δp = total-to-static pressure drop across the machine, inches water
q = heat rate, Btu/hr
t = temperature, F
T = absolute temperature, R (degrees Rankine)
W = air flow rate, pounds/second
y = variable of integration

Paper 57-460, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 1, 1957; made available for printing March 21, 1957.

WALTER ROBINSON is a consulting engineer, Columbus, Ohio.

This paper is based on work performed at The Ohio State University under Air Force sponsorship. Permission granted for this presentation is gratefully acknowledged. Dr. E. O. A. Naumann supervised the work in behalf of the Air Force and made many valuable suggestions. E. O. Buxton, Jr. and J. R. Barnum had charge of the earlier experimental work, and J. R. Mayfield accomplished most of the data correlation and contributed substantially to the study of transient data and damage evaluation.

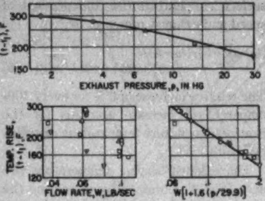


Fig. 1. Data on effect of air pressure on alternator field winding surface temperature and correlation

Z = insulation temperature integral, dimensionless
β = inverted temperature time constant, per hour
η = heat transfer effectiveness, dimensionless
τ = time, hours

Subscripts:

1/2 = condition when half of potential temperature increase has occurred
1 = inlet air
2 = outlet air
∞ = equilibrium condition (asymptotic limit of exponential temperature transient)
B = exhaust bearing
C = commutator
d = heat dissipation rate
ev = equivalent heat capacity
f = condition at end of transient
g = heat generation rate
m = mean (average value)
max = maximum value
x, y, z = conditions at selected times

Correlation of Steady-State Thermal Test Data

Rational methods, notably summarized by Naumann[1] and Friedman,[2] have been found applicable to various blast-cooled machines. Observations relating to the effects of air flow rate, air temperature, exhaust pressure (altitude), load, speed, and machine temperature on pressure drop, component temperature rise, heat transfer effectiveness and heat generation rate of a 3-phase alternator and a d-c generator[3] are discussed.

AIR FLOW RATE

The correlations of pressure drop data for both machines, determined over a wide range of operating conditions, have essentially the same form. The respective equations for the alternator and the generator, based on test data obtained when $[\Delta p/(13.6p)]<0.5$, are

$$W=[0.45/T_m^{0.546}]\times [p\Delta p+0.65\times 10^{-9}\,p^2N^2]^{0.546} \qquad (1)$$

and

$$W=[0.83/T_m^{0.643}]\times [p\Delta p-1.25\times 10^{-9}\,p^2N^2]^{0.530} \qquad (2)$$

It may be noted that for both machines the mean absolute air temperature in the flow passages T_m affects the flow rate. The larger exponent of T_m in equation 2 may indicate that a greater part of the over-all pressure drop occurs near the exhaust end. At a maximum exhaust air temperature on the order of 250 F (121 C), and over the range of probable cooling air temperatures, T_m may vary from 100 to 200 F (38 to 93 C). Therefore, basing flow rate prediction on a constant mean value of T_m would result in $\pm 5.7\%$ maximum error. This error would be greater when either equation 1 or equation 2 was used in long-life or overload rating calculations in which appreciably lower, or respectively higher, exhaust air temperatures would be considered.

The p^2N^2-terms in equations 1 and 2 have opposite signs, indicating the machine rotors' natural tendencies to induce or reduce flow respectively through the machine passages. By means of these correction terms, pressure drops measured at standstill and ground-level exhaust pressure have been correlated with those measured at speeds from 3,000 to 7,000 rpm, and exhaust altitude pressures from zero to 65,000 feet. The correction for fan effect by reference to exhaust pressure correlates experimental data adequately because when p is large, Δp is small, and oppositely. Therefore the actual pressure within the machine at which this effect is created is not discernible from experimental data.

Since the exponents of the corrected pressure drops are greater than 0.5, the predominant effect of the machine passages' frictional characteristics on air flow resistance is indicated. This casts doubt upon the validity of equations 1 and 2 in the range $1>[\Delta p/(13.6p)]>0.5$, which has not been investigated because of limitations in the test apparatus. It is likely that better correlation over the entire range could be achieved by replacing $p\Delta p$ with $[p+(\Delta p/27.2)]\Delta p$.

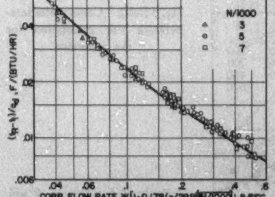


Fig. 2. Heat transfer correlation of d-c generator with pressure and speed correction

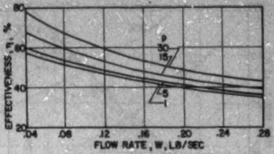


Fig. 3. Heat transfer effectiveness of d-c generator at 6,000 rpm

EXHAUST PRESSURE EFFECT ON COMPONENT TEMPERATURE RISE

The exhaust pressure affects the surface temperature rise of most machine components, particularly those of the rotor. This effect may remain undetected when average temperatures of current-carrying components are determined by resistance measurements, and the heat flow between the components and their heat transfer surfaces is appreciably impeded by intervening insulation, air gaps, etc. This does not imply that air pressure effects may be neglected since hot-spot temperatures may actually occur on the surfaces of rotating components, as was observed on the internally cooled armature of the d-c generator.

The greatest effects were observed on the ends of the field winding surfaces of the alternator. A typical variation of temperature rise with exhaust pressure, at constant load and hot-spot temperature, is shown in Fig. 1. This variation suggests the additive effect of a rotational (tangential) heat transfer component. At low pressure, this component assumes minor importance since the tangential surface velocity is appreciably smaller than the axial air velocity. Thus, an air pressure change of a given percentage has a greater effect, the higher the air pressure. The effective air flow rate defining the heat transfer characteristics of the component may be expressed by

$$W'=W(1+Cp) \qquad (3)$$

where C depends on the component location, its tangential velocity and the local air velocity. The suitability of this type of correction is also shown in Fig. 1 for a flow rate range of 3 to 1 and an exhaust pressure range of 17 to 1. It should be noted that when an attempt is made to define the effective flow rate as $W'=CWp^x$, using ground-level and maximum-altitude test data to determine x, temperature rises at intermediate altitudes may be predicted as much as 10% too low, as shown in Fig. 1. Oppositely, the extrapolation of the data to higher altitudes would result in overestimates of com-

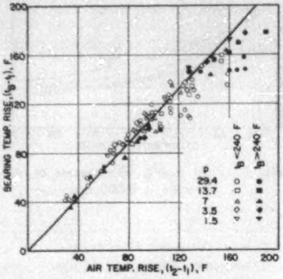


Fig. 4. Exhaust-bearing temperature rise correlation for d-c generator

ponent temperature rise under conditions of usually marginal air flow.

General Heat Transfer Correlation

Friedman has shown that component temperature rise may often be correlated best with component loss.[2] However, if possible it is convenient to define machine heat transfer characteristics on the bases of the average temperature rise of the major heat-dissipating critical component and the total measured heat dissipation of the machine to the cooling air. This correlation can also be used in the analysis of transient thermal conditions, as discussed subsequently, and is believed to be preferable to the use of calculated heat dissipation rates and measured temperature rises of individual components. Temperature rises obtained at various machine temperature levels may be correlated, as shown in Fig. 2, for the d-c generator. The machine temperature t_R is the average of three armature surface temperatures, including the armature hot spot, which occurs at the exhaust end of the lamination stack. The heat dissipation rate q_d is determined from measurements of air temperature rise across the machine. The data in Fig. 2 are for loads from 200 to 500 amperes, exhaust pressures from 1.5 to 29.4 inches Hg, and speeds from 3,000 to 7,000 rpm. The correction of air flow rate for speed of rotation shown in Fig. 2 has not been established reliably because of the relatively small speed range. While it should be expected that $W'=W(1+CpN)$, the smaller exponent on N may account for the greater air temperature increase upstream of the armature laminations, occurring at higher speed. The slope of the correlation line in Fig. 2 is seen to be variable so that a single correlation equation cannot be written. The greater slope at small flow rate may be caused partly by greater air temperature rise upstream and partly by an increase of the part of the total heat dissipation taking place from the armature surfaces, because of increased conduction from the commutator.

A useful form for expressing the heat transfer characteristics of a machine is in terms of its heat transfer effectiveness, defined for the d-c generator as

$$\eta=(t_2-t_1)/(t_R-t_1) \qquad (4)$$

Since

$$(t_2-t_1)=q_d/(868\,W) \qquad (5)$$

and (t_R-t_1) is also directly proportional to q_d, as indicated in Fig. 2, η is independent of load. As shown in Fig. 3, the effectiveness varies little over a small range of air flow rate. Therefore it may be used for rapid estimation of the mean air temperature in the machine passages for a specified machine temperature and an approximately known flow rate.

Component Temperature Rise Correlations

The temperature rise values of critical machine components may be correlated with the rise of the principal correlation temperature or other variables defining the thermal condition of the machine. For example, the hot-spot temperature rise of the d-c generator's armature, occurring at the exhaust end of the lamination stack, is correlated within less than 2% with the armature's average surface temperature rise, independent of speed and exhaust pressure, by

$$(t_{R-\max}-t_1)/(t_R-t_1)=1.07[0.88+0.03(I/100)] \qquad (6)$$

The relative increase of 3% for every 100-ampere load increase accounts for the change of armature temperature distribution with load.

The exhaust-bearing temperature rise data (Fig. 4), obtained by measuring the temperature of the rotating inner race, which is appreciably higher than that of the outer race, exhibit appreciable scatter. Neglecting ground-level high-temperature data where the rise was reduced by heat loss from the frame, a conservative correlation within 10% of the measured values is

$$(t_B-t_1)=1.1(t_2-t_1) \qquad (7)$$

All data points corresponding nearly to the limiting temperature of 250 F (121 C) show a smaller rise than equation 7 gives.

Extensive data of commutator temperature rise obtained on the d-c generator are so inconsistent that a reliable correlation cannot be accomplished. The principal discernible factors influencing commutator temperature, which is essentially the same as brush temperature, are speed of rotation, load, and air flow rate. An approximate correlation equation shown in Fig. 5 to fit the data within about 20%, when the lowest 10% of the points are omitted, is

$$(t_C-t_1)=0.235[0.04(N-3{,}000)+I]/W^{0.474} \qquad (8)$$

The bracketed term defines empirically the variation of the commutator's heat dissipation rate to the cooling air with speed and load. The variation of the commutator's heat generation rate is not necessarily analogous because of conduction to the machine frame and the lamination stack.

Heat Generation Rate

Effects of machine temperature on the heat generation rate (losses) of the machines are discernible from measurements of cooling air temperature rise under con-

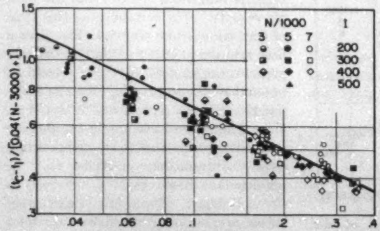
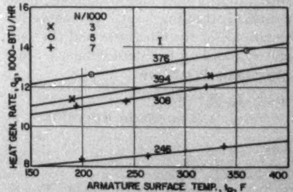


Fig. 5 (left). Commutator temperature rise correlation for d-c generator

Fig. 6 (right). Effect of armature temperature on heat generation rate of d-c generator



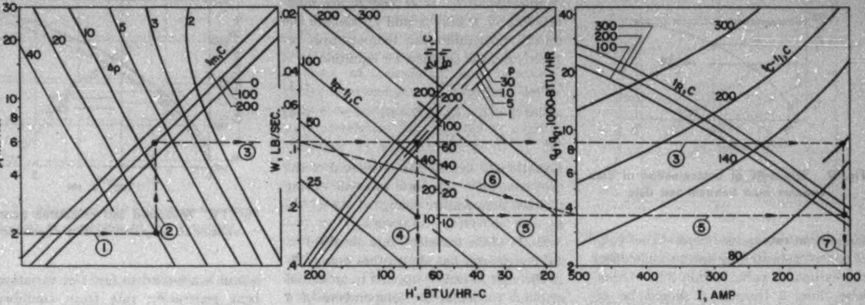


Fig. 7. Chart presentation of thermal characteristics of d-c generator at 6,000 rpm

trolled conditions. Procedures used by Friedman involving insulation of the machine frame and shrouding of the exhaust ports are desirable to obtain accurate data.[4] Typical data of the d-c generator shown in Fig. 6 indicate that at constant load and speed, the heat generation rate may be expressed closely by

$$q_g = G + Bt_R \tag{9}$$

In rating tests, a range of loads and machine temperature conditions would have to be covered, which would be encountered under steady-state and transient operating conditions. Since under temporary overload conditions machine temperatures may be reached which would result in considerable damage or failure if sustained continuously, extrapolation methods must be employed to determine heat generation rates under these conditions.

Graphical Presentation of Machine Thermal Characteristics

The 3-quadrant rating chart proposed by Naumann[1] provides a convenient means for speedy evaluation of permissible loads at specified cooling conditions, based on uniquely defined limiting machine temperatures. It does not include the effect of air temperature increase in the machine passages on flow rate and the effect of exhaust pressure on machine temperature rise. With some additional complexity, a chart may be devised in which these effects are included and which would permit the determination of permissible load on the basis of component temperatures compatible with the machine life required by the specific application. Such a chart for the d-c generator is illustrated in outline in Fig. 7 for one operating speed (6,000 rpm). The left quadrant contains a graphic presentation of the air flow characteristic defined by equation 2, the center quadrant contains a graphic solution of the heat transfer characteristic given in Fig. 2 for the heat dissipation rate q_g, and an alignment chart for determining the cooling air temperature rise (equation 5) and the bearing temperature rise (equation 7). The right quadrant contains a presentation of the machine heat generation characteristics for a range of average armature surface temperatures t_R, and a presentation of the approximate commutator temperature rise correlation (equation 8), using the flow rate scale of the center quadrant as ordinate. The abscissa of the left quadrant is

$$\log[(p\Delta p - 1.25\times 10^{-9}\, p^2N^2)^{0.536}]$$

The heat transfer factor H, shown as the abscissa of the center quadrant, is useful in transient analysis and is defined by

$$H = q_d/(t_R - t_1) \tag{10}$$

Its values are reciprocal to those indicated by the ordinate of Fig. 2. Equal logarithmic ordinate scales for W and q and increasing in the opposite direction, permit locating a logarithmic scale halfway between and with half the length per cycle, which indicates $(t_2 - t_1)$ and $(t_B - t_1)$ at the intersection with a line connecting corresponding values of W and q_d.

For specified values of p, Δp, t_1, and t_R, the permissible load may be determined by means of projection lines drawn in the order of their numbers. The heat transfer effectiveness plot (Fig. 3) gives a close estimate of t_m in the range of possible flow rates, indicated by line 2, since

$$t_m = t_1 + \eta(t_R - t_1)/2 \tag{11}$$

By means of lines 3, 4, and 5, the heat

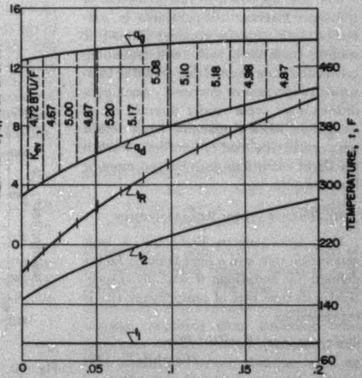


Fig. 8. Graphic determination of equivalent heat capacity of d-c generator from temperature data of overload test

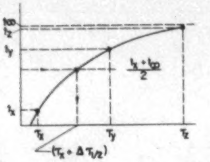


Fig. 9. Schematic of determination of time constant from transient test data

dissipation rate q_d is found. The value of (t_2-t_1) indicated by line 6 should check the estimated value within 3%. Otherwise, lines 3 through 6 should be redrawn, starting with a corrected t_m value. Extending line 5 to the specified t_R line in the right quadrant gives the permissible load. Extending line 3 to the right quadrant and intersecting with the extension of line 7 indicates the commutator rise. Line 6 gives the exhaust bearing rise. Both must correspond to allowable component temperatures. If the commutator temperature is excessive, the intersection of line 3 with the allowable (t_C-t_1) line defines the permissible load. If the bearing temperature is excessive, a new line 6 is drawn through the allowable value of (t_B-t_1) for the same line 3. The intersection of the new line 6 with the ordinate of the right quadrant determines the proper location of a new line 5, which at its intersection with line 4 determines a value of (t_R-t_1) smaller than the allowable. This determines the t_R line to be intersected with the new line 5 in the right quadrant, indicating the permissible load.

Prediction of Transient Machine Temperatures

The first approach to the prediction of transient machine temperatures is concerned with stepwise changes of load or cooling conditions, with both remaining constant during the ensuing period of temperature increase or decrease. Such conditions may occur under aircraft cruise conditions when a machine must accept temporarily the load of another, or when the flight conditions are changed more or less abruptly.

TIME-TEMPERATURE RELATIONSHIPS

If a homogeneous body has a heat generation rate and a heat transfer factor defined by equations 9 and 10 respectively, its time rate of temperature rise is expressed by

$$dt/d\tau=(q_g-q_d)/K=[(G+Bt)-H(t-t_1)]/K \quad (12)$$

where K represents the heat capacity per degree rise of t. If H and K are independent of t, and t_0 and t_∞ denote the initial and equilibrium temperatures respectively, the solution for equation 12 is

$$t=t_\infty-(t_\infty-t_0)e^{-\beta\tau} \quad (13)$$

where $t_\infty=(G-Ht_1)/(H-B)$ (14)

and $\beta=(H-B)/K$ (15)

Equation 15 defines the reciprocal of the time constant, the time interval during which a temperature change of $[1-(1/e)](t_\infty-t_0)=0.632(t_\infty-t_0)$ occurs.

Study of the transient data obtained on a d-c generator has shown that equations 12 through 15 can be applied to an actual machine to define the temperature of a critical machine component, such as the armature, which is representative of a major part of the machine's heat-generating and heat-storing mass. The heat capacity is then an equivalent value which associates the temperature change of that component with the total heat storage, generation, and dissipation abilities of the machine. Therefore the steady-state characteristics defining G, B, and H in relation to that correlation temperature can be employed in transient analysis. The heat capacity must be determined from transient test data.

DETERMINATION OF HEAT CAPACITY

In transient tests at constant air flow rate, rates of machine temperature increase may be produced that are comparable to those encountered under abruptly applied overload conditions. These tests yield data on variation of machine temperature and exhaust air temperature with time from which the equivalent heat capacity may be calculated by

$$K_{ev}=\left[\int_{\tau_x}^{\tau_y}(q_g-q_d)d\tau\right]/(t_{R-y}-t_{R-x}) \quad (16)$$

providing that all heat dissipation occurs to the cooling air. As shown in Fig. 8, the test data on t_R (which may be the average of more than one temperature), t_1 and t_2 are used to find the variation of heat generation rate from equilibrium test data of q_g versus t_R (Fig. 6) and the variation of heat dissipation rate from equation 5. Graphic integration of the area between the curves for q_g and q_d for arbitrarily chosen increments of t_R yields values of K_{ev} according to equation 16, as shown in Fig. 8. This method has been found to give consistent values[5] and has the advantage that the equivalent heat capacity is being evaluated on the basis of data for which the temperature distribution in the machine is similar to that under overload conditions.

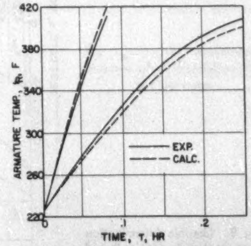


Fig. 10. Measured and calculated transient armature temperatures of d-c generator

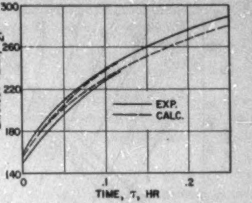


Fig. 11. Measured and calculated transient exhaust air temperatures of d-c generator

Another method makes use of equation 15 and requires the determination of β from transient test data. It can be shown that

$$\beta=0.693/\Delta\tau_{1/2} \quad (17)$$

where $\Delta\tau_{1/2}$ is the time interval during which half of the rise from any temperature t_x to t_∞ would take place. If the transient test is not run to equilibrium, the value of t_∞ may be calculated on the basis of the temperatures measured at beginning and end of two equal and successive time intervals $(\tau_y-\tau_x)=(\tau_z-\tau_y)$, shown schematically in Fig. 9, by

$$t_\infty=t_z+\{(t_z-t_y)^2/[(t_y-t_x)-(t_z-t_y)]\} \quad (18)$$

Then the value of $\Delta\tau_{1/2}$ is defined by the location of $(t_x+t_\infty)/2$ on the curve, as shown in Fig. 9. This method has not been applied satisfactorily to short-transient data.[5] It has been found to give consistent values of K_{ev} only when the test data extend nearly to equilibrium. However, the values so determined for a d-c generator[5] were 8% lower than those based on short-transient data and the graphical method shown. The principal reason for this discrepancy is that the different internal machine temperature distribution which develops during long transient tests does not correspond to that encountered under overload conditions. Therefore the graphic method should be considered preferable.

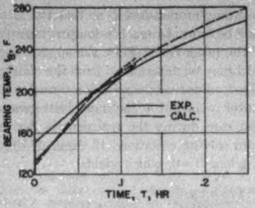


Fig. 12. Measured and calculated transient-bearing temperatures of d-c generator

PREDICTION OF MACHINE TEMPERATURES

If, as is assumed in the analysis, q_g varies linearly with the principal correlation temperature t_R, the values of G and B may be defined directly for any load condition. Since H depends on the cooling conditions and K_{ev} is a machine constant, t_R at any time may be calculated with the use of equations 15, 14, and 13. If q_g versus t_R is not linear, B may be approximated by

$$B=(q_{g-f}-q_{g-0})/(t_{R-f}-t_{R-0}) \tag{19}$$

The final temperature t_{R-f}, must be assumed in order to estimate the final heat generation rate q_{g-f}, and must then be verified by equations 15, 14, and 13. Fig. 10 shows comparisons of calculated and measured armature temperatures of a d-c generator,[5] which are typical of the maximum deviations observed among the overload test data.

When the final temperature is specified, the time interval required to attain it may be calculated from

$$(\tau_f-\tau_0)=(1/\beta)\log e[(t_\infty-t_0)/(t_\infty-t_{R-f})] \tag{20}$$

with the use of equations 14, 15, and, if necessary, 19.

Regardless of the variation of q_g versus t_R, the chart of Fig. 7 may be used to find the values of H, G, and B for a transient of known load, initial machine temperature, and approximate final machine temperature. The value of B may be determined from equation 19, with quantities found in the right quadrant. Then

$$G=q_{g-0}-Bt_0 \tag{21}$$

If the values of Δp and p are specified so that the flow rate may change during the transient because of increase in mean air temperature in the machine, the value of t_m to be used in finding H from the chart should be conservatively estimated by $(t_1+t_{R-f})/2$.

The transient test data of a d-c generator[5] have shown that the cooling air temperature rise during any part of the transient is defined by equation 5, using the steady-state heat transfer correlation of Fig. 2 to determine q_d corresponding to the instantaneous value of (t_R-t_1). Comparisons of calculated and measured values representative of maximum deviations are shown in Fig. 11. The steady-state correlation for the exhaust bearing (equation 7) was not found to fit the transient data. For the bearing, and the commutator as well, empirical correlations were developed which gave reasonably good predictions. The time rate of exhaust air temperature increase was used as a correlation parameter for the bearing. The commutator correlation was based on a parameter proportional to the initial ratio of heat transfer rates between the commutator and the armature by conduction and between the commutator and the cooling air by convection, as

$$(t_{C-0}-t_{R-0})/[(t_{C-0}-t)W^{0.464}]$$

Comparisons of bearing and commutator temperatures, calculated on the bases of these correlations, with values measured in overload tests are shown in Figs. 12 and 13 respectively. They are representative of maximum differences between prediction and measurement under test conditions resulting in relatively high temperatures of these components.

Evaluation of Cumulative Insulation Damage from Transient Thermal Conditions

The prediction of temperatures during transient thermal conditions is important for the determination of potential catastrophic failure of a component or for the evaluation of the permanent damage which may be caused during this operating interval. The accumulation of permanent damage is particularly pertinent to electrical insulating materials for which life versus temperature data are applicable to transient temperature conditions, since the rate of damage accumulation, if considered a function of temperature only, would be at any temperature reciprocal to the insulation life. Thus

$$dD/d\tau=f(T)=1/L \tag{22}$$

When failure finally occurs, $D=1=\Sigma\Delta D$, representing the sum of damage increments ΔD, which have been accumulated during periods of steady-state and transient operation at various temperature levels. The damage increment for each operating period from τ_0 to τ_f is

$$\Delta D=\int_{\tau_0}^{\tau_f}(1/L)d\tau \tag{23}$$

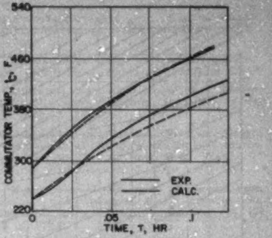


Fig. 13. Measured and calculated transient commutator temperatures of d-c generator

If the temperature is constant, L does not vary with time and $\Delta D=\Delta\tau/L$. For a temperature transient which follows equations 12 through 15, so that

$$dT/d\tau=\beta(T_\infty-T) \tag{24}$$

equation 23 may be transformed to

$$\Delta D=(1/\beta)\int_{T_0}^{T_f}dT/[L(T_\infty-T)] \tag{25}$$

When the life-temperature relationship is defined by a plot of log (L) versus T, as

$$L=Ce^{-aT} \tag{26}$$

where C and a are characteristic constants of the insulation,[9] equation 25 is evaluated by

$$\Delta D=[e^{aT_\infty}/(\beta C)]\{-Ei(-x_f)-[-Ei(-x_0)]\} \tag{27}$$

$$\text{where } x=a(T_\infty-T) \tag{28}$$

$$\text{and } -Ei(-x)=\int_x^\infty(e^{-y}/y)dy \tag{29}$$

Values of the function $-Ei(-x)$ are available in mathematical tables.[10] When $x>15$, the function may be approximated by (e^{-x}/x).

When the life-temperature relationship is defined by a plot of $\log(L)$ versus $(1/T)$, according to the Arrhenius law of reaction rates, consistent with the concept that the deterioration of insulation materials is the cumulative result of chemical and related processes,[11,12] as

$$L=Ae^{b/T} \tag{30}$$

where A and b are characteristic constants, equation 25 is evaluated by

$$\begin{aligned}\Delta D=&[1/(\beta A)]\{-Ei(-w_f)-\\&e^{-s}[-Ei(-(w_f-s))]\}-\\&\{-Ei(-w_0)-e^{-s}[-Ei(-(w_0-s))]\}\\=&[1/(\beta A)][Z(w_f,s)-Z(w_0,s)]\end{aligned} \tag{31}$$

where $w=b/T$ and $s=b/T_\infty$. A chart of the dimensionless integral function Z is shown in Fig. 14. The chart is generally applicable to the evaluation of cumulative damage, resulting from transient temperatures occurring under constant-air-flow conditions, to any machine with an insulation life defined by equation 30.

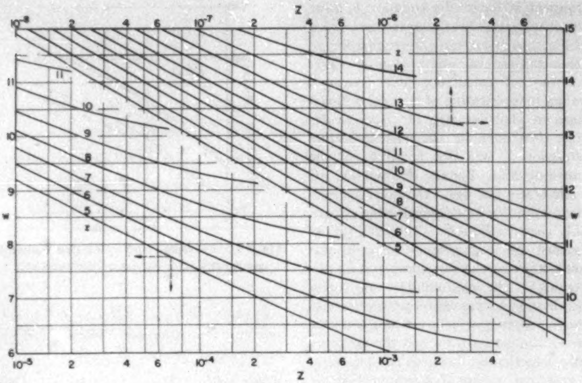


Fig. 14. Chart of integral function for evaluation of cumulative insulation damage

Machine-Use Evaluation and Rating on Basis of Predicted Damage

On the basis of life-temperature data of the machine insulation, a continuous operating temperature may be selected which would assure a specified minimum life. The chart of Fig. 7 would disclose for any cooling condition what the allowable continuous load would be at which this operating temperature occurred. While this type of information is a useful guide in the selection of a machine for a given aircraft application, it will not lead to optimization of machine size and cooling drag since the only safe procedure is to select the machine and its cooling system so that for any expected combination of flight conditions and electrical load the rated machine temperature will not be exceeded. Most of the time, the machine will operate at substantially lower temperature. The insulation life will be several times greater than at the rated maximum temperature. Obviously, the machine could withstand periods of greater temperature and still meet the total life requirements. During such periods, temperatures must not be so high as to cause catastrophic failure of any component.

While for any air flow condition an arbitrary steady-state rating can be assigned to a machine, which may readily be interpreted since it is based on a specific temperature and a total operating life, a meaningful transient load rating is more difficult to assign because the initial machine temperature must be known in order to arrive at a prediction of the final temperature reached and the damage incurred during this overload period. Therefore it would appear that best use of the methods of transient-temperature and cumulative damage analysis can be made in evaluating specific machine applications for maximum component temperatures and damage increments incurred during typical application cycles, including probable emergency-load conditions. In such evaluations, combinations of load and cooling conditions would be specified. For a specific machine, the study of temperature variation and cumulative damage during a typical flight cycle, resulting from use of different air induction systems, should lead to the selection of the best system which represents a compromise between machine life and cooling penalty.

PERMISSIBLE LOAD FOR SPECIFIED DAMAGE INCREMENT

For emergency overload rating of machines, a normal operating temperature prior to overload may be specified. A damage increment may be selected which could be tolerated for a number of times without seriously reducing the normal machine life. For this damage increment different loads would be associated with different durations of overload application. A first estimate of permissible load during a time interval $\Delta\tau$ for a damage increment ΔD may be obtained by assuming that the average damage rate during the transient is equal to the arithmetic mean of the initial and final rates. Thus

$$\Delta D=(\Delta\tau/2)[(1/L)_0+(1/L)_f] \qquad (32)$$

Solving equation 32 for L_f gives an estimate of the life at the approximate final machine temperature t_f so that the latter may be found from a life-temperature plot of the insulation. Then the approximate load may be determined from the chart of Fig. 7. For that purpose H is found in the center quadrant. The mean heat generation rate during the transient is derived from solving equations 13 through 15 for G, when $B=0$, which yields

$$G=(q_{\theta-0}+q_{\theta-f})/2$$
$$=\mathrm{H}\{[(t_f-t_0e^{-H\tau_f/K_{ev}})/(1-e^{H\tau_f/K_{ev}})]-t_1\} \qquad (33)$$

Line 5 in the right quadrant for $q_\theta=G$ indicates loads corresponding to t_{R-0} and the approximate t_{R-f}. The approximate permissible load for the transient is the average of these two values. On the basis of this load, a more accurate evaluation of the final machine temperature and the corresponding damage increment can be made. To meet the ΔD specification, a second approximation may be necessary. For short time intervals during which the machine temperature increase is on the order of 50 C, the approximation of equation 32 yields results which may be adequately accurate for preliminary evaluation purposes.[5]

Conclusions

Effects of air temperature in machine passages on pressure drop and of exhaust pressure on machine temperature rise should be included in predicting machine component temperatures. Use of the machine's heat transfer effectiveness, relating the air temperature rise to a principal correlation temperature rise, facilitates temperature prediction. A relatively simple graphic presentation of machine characteristics can be made which may be used to predict steady-state component temperatures over a wide range of conditions. The chart also aids in the calculation of transient component temperatures.

For transient temperature analysis, an equivalent machine heat capacity can be determined which should be based on the principal correlation temperatures recorded during overload transient tests. From such data, empirical correlations may also be obtained for the prediction of transient temperatures of other critical components. The heat transfer process to the cooling air during transient conditions is the same as in the steady-state and can therefore be evaluated from steady-state data. Calculations for transient temperature prediction are not complicated, but may be deemed laborious.

Using constant-temperature insulation life data and defined transient thermal characteristics, cumulative insulation damage evaluations can be made. Such evaluations of load and cooling cycles of specific applications would give a proper basis for the selection of a machine and of its air induction system.

References

1. CORRELATION OF RATING DATA FOR ROTATING ELECTRIC MACHINERY. **Erwin O. A. Naumann.** *AIEE Transactions*, vol. 73, pt. II, 1954 (Jan. 1955 section), pp. 443–52.

2. CORRELATION OF THE THERMODYNAMIC AND ELECTRICAL CHARACTERISTICS OF BLAST-COOLED GENERATORS, **Daniel Friedman.** *Ibid.*, vol. 74, pt. II, 1955 (Jan. 1956 section), pp. 448–56.

3. THERMAL STUDY AND RATING OF BLAST-COOLED AIRCRAFT ELECTRICAL GENERATORS, **W. Robinson.** *WADC Technical Note 56-121, ASTIA Document No. 110416*, Armed Services Technical Information Agency, Wright Air Development Center, Dayton, Ohio, Apr., 1956.

4. CORRELATION OF THE THERMODYNAMIC AND ELECTRICAL CHARACTERISTICS OF BLAST-COOLED AIRCRAFT GENERATORS, PART II–EXPERIMENTAL EVALUATION OF A GENERAL ELECTRIC 15-KVA NARROW-SPEED-RANGE GENERATOR, **D. Friedman.** *NRL Report 4610*, Naval Research Laboratory, Washington, D. C., Oct., 1955.

5. PREDICTION OF TRANSIENT COMPONENT TEMPERATURES AND THEIR EFFECTS ON THE LIFE OF BLAST-COOLED ELECTRICAL MACHINES, **W. Robinson, J. R. Mayfield.** *WADC Technical Note 56-122, ASTIA Document No. 110417*, Wright Air Development Center, Dayton, Ohio, Apr., 1956.

6. SYMPOSIUM ON TEMPERATURE STABILITY OF ELECTRICAL INSULATING MATERIALS. *ASTM Special Technical Publication No. 161*, American Society for Testing Materials, Philadelphia, Pa., 1954.

7. TEST CODE FOR EVALUATION OF SYSTEMS OF INSULATING MATERIALS FOR RANDOM-WOUND ELECTRIC MACHINERY. *AIEE Standard no. 1C* Jan. 1954.

8. THERMAL EVALUATION OF INSULATING MATERIALS. *AIEE Special Publication S61*, June 1954.

9. ELECTRICAL INSULATION (book), **G. L. Moses.** McGraw-Hill Book Company, Inc., New York, N. Y., 1951.

10. TABLES OF SINE, COSINE, AND EXPONENTIAL INTEGRALS, VOL. II (book), **Works Projects Administration of the City of New York.** Bureau of Standards, Washington, D. C., 1940.

11. ELECTRICAL INSULATION DETERIORATION TREATED AS A CHEMICAL RATE PHENOMENON, **Thomas W. Dakin.** *AIEE Transactions*, vol. 67, 1948, pp. 113–22.

12. GUIDING PRINCIPLES IN THE THERMAL EVALUATION OF ELECTRICAL INSULATING MATERIALS, **L. J. Berberich, T. W. Dakin.** *Ibid.*, vol. 75, pt. III, Aug. 1956, pp. 752–61.

Surface Heat-Transfer Coefficients of Salient Poles in a Blast-Cooled Alternator

W. ROBINSON
NONMEMBER AIEE

F. S. TSE
NONMEMBER AIEE

THERE are few data available on which the calculation of heat-transfer coefficients of blast-cooled rotating machine structures and adjacent stationary internal component surfaces can be based. The application of fundamental heat-transfer relationships, derived from stationary simple configurations, is impractical because of the significant effects of rotation and their variations with density of the cooling air, which may occur in machine passages. From temperature measurements on actual machines, heat-transfer coefficients cannot be derived reliably because of the usual difficulties in determining the heat-flow distribution among the various surfaces of a component whose heat dissipation may be known.

This paper has a twofold purpose: first, to present a part of the data obtained in a comprehensive experimental study of one configuration, namely, the field structure of a salient-pole alternator,[1] and to illustrate thereby typical effects of speed of rotation and of air density on magnitudes and distributions of local heat-transfer coefficients; second, to describe an experimental technique which may be used to gather heat-transfer data on all types of blast-cooled rotating and stationary component configurations employed in the design of aircraft electric machines.

Experimental Method

The principle of the method described in this paper has been applied by several investigators employing different techniques to the study of forced convection heat transfer from stationary and rotating surfaces of simple geometry.[2–6] The method consists mainly in measuring the temperature distribution of a very thin electrically heated resistance ribbon, which is cemented to an electrically insulated surface exposed to air flow. If the configuration of the surface is the same as that of a heat-dissipating machine component, data for the calculation of local heat-transfer coefficients even on very irregularly shaped surfaces can so be obtained quite simply.

NOMENCLATURE

b = thickness of heating ribbon, feet (ft)
F_a = radiation configuration factor, dimensionless
F_e = radiation emissivity factor, dimensionless
h = local heat-transfer coefficient, Btu per hour–square feet–degrees Fahrenheit (Btu/hr–sq ft–F)(=273 watt/sq inch – degrees centigrade)
H = average heat-transfer coefficient, Btu/hr–sq ft–F
I = current, amperes
k = thermal conductivity, Btu/hr–sq ft-(F/ft)
N = speed of rotation, rpm
p = absolute air pressure at test section inlet, inches mercury (Hg)
q = heat rate, Btu/hr
Q = heat flux, Btu/hr–sq ft
R = resistance, ohms/ft
T = absolute local temperature, degrees Rankine (R)
V = velocity, ft/second
w = effective width of heating ribbon, ft
W = air-flow rate, pound/second
X = length, ft
ρ = air density at test section inlet, lb/cubic ft
σ = Stefan-Boltzmann constant, Btu/hr–sq ft–R^4

Subscripts

a = axial flow component, or air
c = convection heat transfer
cd = conduction heat transfer
f = pole face
g = heat generation
m = mean value
o = standstill
r = radiation heat transfer
t = tangential component, or surface
w = environment, or internal housing surface

DETERMINATION OF LOCAL HEAT-TRANSFER COEFFICIENTS

Local heat-transfer coefficients are calculated by making heat balances at points of the resistance ribbon where its surface temperatures are measured. As shown in Fig. 1, heat is conducted along the ribbon itself, in and out of element dx, at rates q_a and q_b. The conduction

---

Paper 57-461, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 5, 1957; made available for printing March 21, 1957.

W. ROBINSON is a consulting engineer, Columbus, Ohio, and F. S. TSE is with Michigan State University, Lansing, Mich.

This paper is based on work performed at The Ohio State University under Air Force sponsorship. Permission granted for this presentation is gratefully acknowledged.

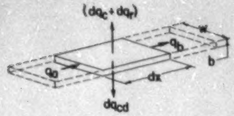


Fig. 1. Schematic of electrically heated resistance ribbon

rate to the mounting surface is dq_{cd}, the radiation rate to the environment is dq_r, and the convection rate to the air flow passing over one ribbon surface is dq_c. The heat-generation rate in element dx is dq_g. By heat balance,

$$dq_c = dq_g + (q_a - q_b) - dq_r - dq_{cd} \qquad (1)$$

Defining each of these heat rates in terms of measurable quantities,

$$h_c w(T - T_a)dx = 3.413 I_m{}^2 R dx + kwb(d^2T/dx^2)dx - \sigma F_e F_a w(T^4 - T_w{}^4)dx - Q_{cd} w dx \qquad (2)$$

Solving equation 2 for the local heat-transfer coefficient yields

$$h_c = [(3.413/w) I_m{}^2 R + kb(d^2T/dx^2) - \sigma F_e F_a (T^4 - T_w{}^4) - Q_{cd}]/(T - T_a) \qquad (3)$$

When a constantan ribbon is used, which is 0.002 inch thick and 3/16 inch wide, and has a resistance of 0.642 ohm/ft and a thermal conductivity of 13.6 Btu/hr–sq ft–F, and the configuration is such that $F_e F_a = 0.7$ (see test apparatus and models in Figs. 2–4), the heat-conduction rate along the ribbon is negligible and equation 3 reduces to the working form

$$h_c = \{140.5 I_m{}^2 - 0.121[(T/100)^4 - (T_w/100)^4] - Q_{cd}\}/(T - T_a) \qquad (4)$$

The local conduction heat flux must be determined from calibration data which relate Q_{cd} with the temperature difference between the resistance ribbon and points within the body of the test model. In irregularly shaped configurations which have surfaces through which appreciable heat losses may occur, guard heaters must be used to supply these losses and to minimize Q_{cd} by virtue of balancing, as best possible, the temperatures of the ribbon surface and the model interior. The calibration data are obtained while the guard heaters are de-energized and the surface covered with resistance ribbons are well insulated. The conduction heat flux is the electrical input into the resistance ribbons, corrected for losses through the insulation. The losses are calculated on the basis of measured temperature gradients in the insulation.

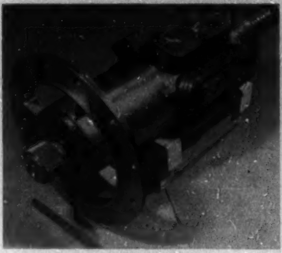


Fig. 3. Rotor assembly with instrumented pole periphery

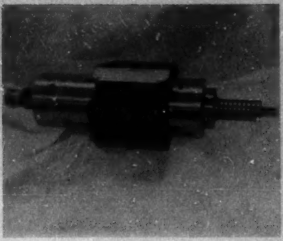


Fig. 4. Rotor assembly with instrumented pole face

Experimental Technique

The experimental study was principally concerned with the measurement of temperature distributions on a rotating model simulating the configuration of a blast-cooled 6-pole 400-cycle 3-phase alternator, which should have a rating of 40 kva, 115/208 volts, at 8,000 rpm. The test apparatus was so designed as to be also applicable to the heat-transfer study of other blast-cooled rotating configurations.

TEST APPARATUS

The apparatus consisted of a system of stationary air ducting, a test section containing the test rotor, and an external variable-speed drive. The principal parts of the apparatus are shown in Fig. 2. At ground-level pressure, air flow was supplied by a blower and metered with sharp-edge orifices. At reduced pressure, air flow was induced by means of a 3-stage steam ejector connected to the vertical discharge pipe, and was metered by means of calibrated inlet nozzles operating at subcritical discharge pressures, from 14 to 1.5 inches Hg absolute. As shown in Fig. 2, the air flow approached the test rotor through an annular passage, formed by the thermocouple switch unit and its housing, and having the same radial dimension as the pole models mounted to the rotor. The speed of the rotor could be varied from 1,000 to 9,600 rpm. The rotor is shown removed from the apparatus in Figs. 3 and 4. The slip rings used to supply heating current to the test models are apparent at the right of Fig. 4. The connector for the thermocouple leads between the test model and the switch unit is seen in the foreground of Fig. 3.

TEST CONFIGURATION AND MODELS

The proportions of the simulated pole pieces are apparent in Figs. 3 and 4. Their dimensions are given in Fig. 5. They were all made of fiber-base phenolic material. Two pole models were instrumented for surface heating and temperature measurements, but were used separately in the tests. On one pole model, heating ribbons were cemented to the periphery (Fig. 3). On the other pole

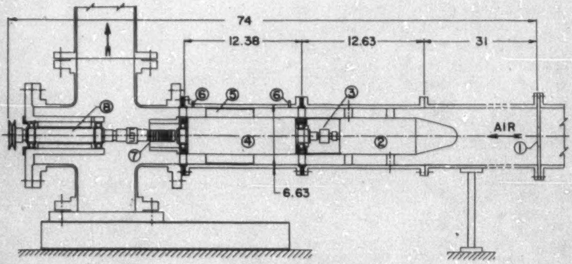


Fig. 2. Test apparatus

1—Metering orifice
2—Thermocouple switch
3—Thermocouple connector
4—Rotor
5—Poles
6—Air-temperature probes (stations 2–3)
7—Slip rings
8—Drive shaft

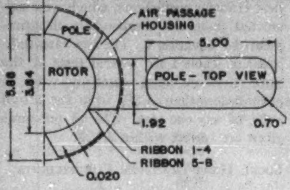


Fig. 5. Dimensions of rotor test model

model, the ribbons were cemented to the pole face (Fig. 4). By rotating in both directions, these two models could be used to define the entire surface heat transfer distribution of one pole. All heating ribbons were of constantan, 0.002 inch thick and 3/16 inch wide, and were connected in series. Guard heaters were made of the same material, as shown on the pole face in Fig. 3 and the pole periphery in Fig. 4, and were either connected in series with the heating ribbons, or were separately energized and controlled.

The constantan ribbons were moderately etched and bonded to the pole models with epoxy cement. Four ribbons instrumented for temperature measurements (*1–4*, in Fig. 5) were placed on the side surfaces to indicate not only the peripheral variation, but also the radial variation in local heat transfer coefficients. Several instrumented ribbons were also placed on the pole face, but mainly the one near the edge (*5*, in Fig. 5) was used to determine the variations of local heat-transfer coefficients in the axial and tangential directions.

TEMPERATURE MEASUREMENTS

Three types of iron-constantan thermocouples were installed in the instrumented pole models, as indicated schematically in Fig. 6 (*A*, *B*, *C*). Thermocouples for measurement of ribbon temperatures were formed by spot welding 30-gage iron wires to the undersides of the constantan ribbons (*A*). Other thermocouples were formed of 30-gage iron and constantan wires (*C*), and placed in the phenolic material to provide information on internal temperature gradients from which conduction heat losses could be determined on the basis of the thermal calibration data of the test configuration. The third type were differential thermocouples (*B*) which were placed with their junctions 1/16 and 7/16 inch beneath the heated surfaces so as to provide information for the adjustment of the guard heaters' input which was desired to produce, if possible, null readings of these differential thermocouples. The model shown in Fig. 3 had 30 surface and 13 embedded thermocouples. The model in Fig. 4 had 14 surface and 5 differential thermocouples.

As shown in Fig. 6, a common constantan lead was used for all thermocouples and all temperature measurements were referred to a reference junction located in the rotor. Therefore, all thermocouples in the pole models were of the differential type, which reduced the switch requirement for measurements with each to connecting only one iron lead. The temperature of the reference junction was measured accurately by means of an aged thermistor (0.040-inch ceramic bead), which was calibrated against the reference thermocouple at standstill. During rotation, the reference thermocouple, being connected, like all other thermocouples, through silver-carbon slip rings to a precision potentiometer, would give erroneous readings because of brush friction, while the thermistor circuit, because of its high resistance (8×10^4 ohm), was insensitive to brush contact resistance, as long as the brushes were maintained in reasonably good condition.

Details of the thermocouple switch, which was required to transmit readings of 45 rotating thermocouples to a stationary potentiometer (see Figs. 2 and 6), are shown in Fig. 7. Its rotor assembly consisted of a 48-point connector, a 12-position 4-wafer selector switch actuated by a rotary solenoid, and 5 silver slip rings. The switch was so connected that in each position the temperature differences between four points and the reference junction could be read. The common iron lead (Fig. 6) was connected directly to one of the slip rings and the constantan lead from the reference junction passed directly to the button at the end of the slip ring shaft (7, Fig. 7). In one position of the selector switch, the common iron lead was shorted to the iron leads from the four other slip rings. Any readings obtained in this position indicated errors due to temperature differences in the switch circuit and were used to correct the data.

The constantan ribbons were part of the heating and thermocouple circuits and 400-cycle a-c power was used for heating. A capacitor (Fig. 6) was placed in the power circuit to block the thermo-

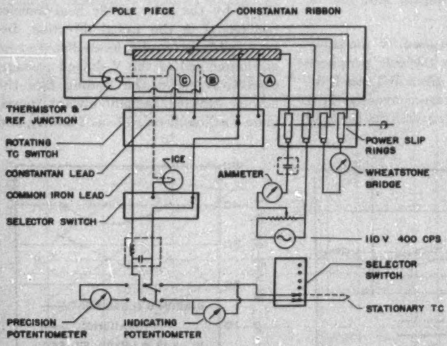


Fig. 6. Schematic of thermocouple and heating circuits

A—Surface thermocouple
B—Differential thermocouple
C—Body thermocouple

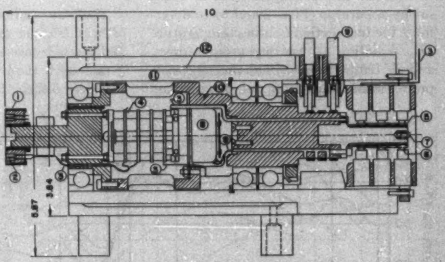


Fig. 7. Rotating thermocouple switch unit

1—Thermocouple connector
2—Thermocouple contact
3—Thermocouple lead
4—Circuit selector
5—Silver slip ring
6—Silver-carbon brush
7—Constantan button
8—Solenoid
9—Power and signal leads
10—Thermocouple switch rotor
11—Stationary housing
12—Water jacket

Table I. Test Conditions and Average Heat-Transfer Coefficients

Test Conditions		Average Heat-Transfer Coefficient, Btu/Hr-Sq Ft-F Speed of Rotation, Rpm						
Air Rate, Lb/Sec	Pressure, In. Hg	0	1,200	1,300	2,400	4,800	6,400	8,000
				Pole Periphery				
0.024	29.4	4.0	9.6		13.2	21.4		
0.049	29.3	7.1				23.4	28.0	30.0
0.093	29.2	9.9	11.9		14.8	22.9	27.5	35.2
0.147	28.8					23.5	27.2	35.2
0.200	29.2	15.2				22.2		
0.300	29.5	23.6				30.0		
0.399	29.3	31.6	30.7		30.8	33.3	35.0	37.0
0.501	29.7	32.7				39.0		
0.038	13.8							18.6
0.038	7.0							12.0
0.038	3.5	5.5						7.7
0.038	1.5							6.5
0.063	13.7							19.0
0.063	7.0							12.3
0.063	3.4	8.4						9.3
0.109	13.6	11.5				14.1		18.8
0.109	7.1	12.3				12.8		13.7
0.109	3.7	10.6	11.0		11.4	12.3		12.2
0.175	5.9	17.5						
0.176	7.0							17.9
0.258	8.2							20.5
				Pole Face				
0.026	28.9	7.8				22.8		31.8
0.005	29.4	9.4				23.7		32.0
0.114	28.4			12.9	15.7			
0.317	29.1	18.9				25.4		33.6
0.039	12.1							18.6
0.039	3.4	8.6				8.7		10.8
0.039	2.0							9.2
0.039	1.6	8.6						
0.063	4.2					10.2		11.8
0.063	3.3	9.2						
0.110	14.0							16.3
0.109	4.8					11.2		10.6
0.109	3.7	10.7		11.1	10.7			
0.283	14.2	16.8						20.5
0.282	8.2	16.6				15.7		16.0

couple electromotive force. A filter choke (Fig. 6) was introduced in the thermocouple circuit so that the alternating voltage across the ribbons would not be imposed on the potentiometer.

A number of stationary 30-gage iron-constantan thermocouples were installed in the apparatus (Figs. 2 and 6) to determine the test section housing temperature (T_w, equation 4), bearing temperatures, and air temperatures. The thermocouple probes located upstream and downstream from the test model (stations *2* and *3*, Fig. 2) were shielded and designed for traversing the flow passage.

Experimental Results

RANGES OF DATA

The test conditions used for the study of heat transfer from the pole periphery and the pole face are given in Table I, together with the average convective heat transfer coefficients derived from the local heat transfer coefficients, which were calculated from the temperature measurements by equation 4. The flow rates indicated in Table I for several tests at the same speed are mean values and differ, in general, less than ±3% from the actual value of any one test. The air pressures given are correct within ±1%.

LOCAL HEAT-TRANSFER COEFFICIENTS

In calculating the local heat-transfer coefficients by equation 4, it was assumed that the air temperature increased linearly between the two measuring stations indicated in Fig. 2. The error introduced by this assumption was small since the maximum temperature rise of the heating ribbons was less than 40 F and most of the air temperature rise was due to windage. Typical distributions of local coefficients on the pole periphery and face are shown in Figs. 8 and 9, respectively.

The left halves of the plots in Fig. 8 are for the leading side of the pole, the right halves are for the trailing side. The average heat-transfer coefficient obtained by graphical integration of the curves in Fig. 8 is 35.2 Btu/hr-sq ft-F (Table I). Thus, typically, maximum and minimum local coefficients differ from 35 to 60% from the average value and occur at the leading and trailing corners of the pole's upstream (front) end. The points at the front and rear center lines of the poles have about the same coefficients, but the variation in the radial direction is much greater in the rear. The occurrences of maximum and minimum values may be associated with flow impingement and separation, respectively. The circulatory nature of the flow in the V-shaped passages between the poles is indicated in Fig. 8 by the order of the heat-transfer coefficients in the radial direction, between ribbons *1* and *4*, located at the root and outer end of the V-shaped passage, respectively. On the leading face the flow is directed radially inward (maximum coefficients on ribbons *3* and *4*), and

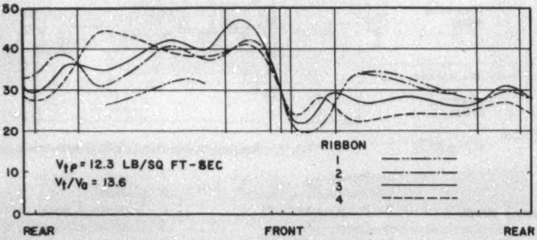


Fig. 8. Typical distribution of local heat-transfer coefficients on pole periphery

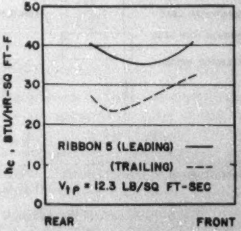
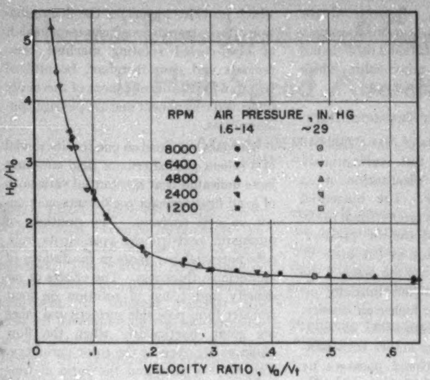


Fig. 9. Typical distribution of local heat-transfer coefficients on pole face



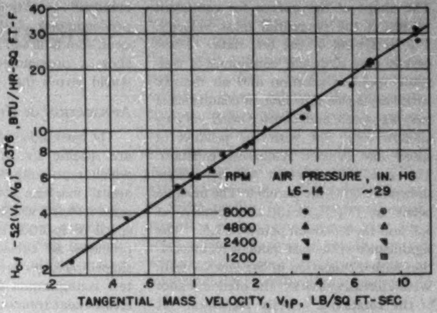


Fig. 10 (left). Correlation of average heat-transfer coefficients of pole periphery

Fig. 12 (above). Correlation of average heat-transfer coefficients of pole face

on the trailing face it is directed outward (maximum coefficients on ribbons *1* and *2*).

The average heat-transfer coefficient for the pole-face data shown in Fig. 9 is 32.0 Btu/hr-sq ft–F (Table I). Thus, typically, maximum and minimum local coefficients differ from 15 to 30% from the average value. The rotation is shown to increase heat transfer at the leading half of the pole face. With large flow rates or at small air pressure the difference between the leading and trailing halves becomes less pronounced. The axial variation of local coefficients decreases with the flow rate. The air-flow pattern in the 0.020-inch gap between the pole face and the housing is principally determined by the tangential velocity of the rotating face, and is modified by the flow rate to a small extent, but more so when the air pressure is low.

Correlations of Average Heat-Transfer Coefficients

The data of Table I show that heat-transfer coefficients increase with flow rate, speed of rotation, and air pressure. The effect of flow rate at constant speed is greater, the smaller the air pressure. Similarly, at constant flow rate, the effect of speed is greater, the greater the air pressure. At the same air pressure, the effect of speed is greater, the smaller the flow rate.

POLE SIDE SURFACES

The average heat-transfer coefficients of the pole periphery at standstill are correlated by the following equation

$$H_o = 56.73W^{0.71} = 11.7(V_a\rho)^{0.71} \qquad (5)$$

The axial air velocity, V_a, is that in the annular approach to the rotating configuration. As shown in Fig. 10, the heat-transfer coefficients determined with rotation over the entire range of air pressures are correlated nondimensionally with the standstill coefficient by the ratio of the axial air velocity to the mean tangential velocity of the pole side surfaces. The correlation equation for the range $0<(V_t/V_a)<26$ is

$$H_c/H_o = 1 + 0.0316(V_t/V_a)^{1.71} - 48.5e^{-60.9(V_a/V_t)} \qquad (6)$$

The presentation of equations 4 and 5 in Fig. 11 is convenient for the direct determination of heat-transfer coefficients.

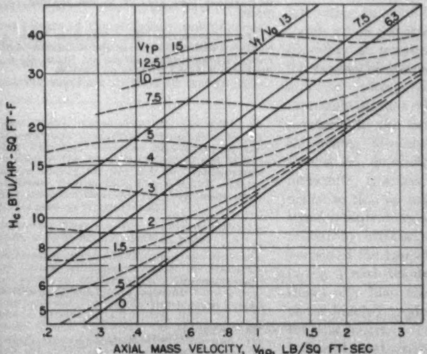


Fig. 11. Chart of average heat-transfer coefficients on pole periphery

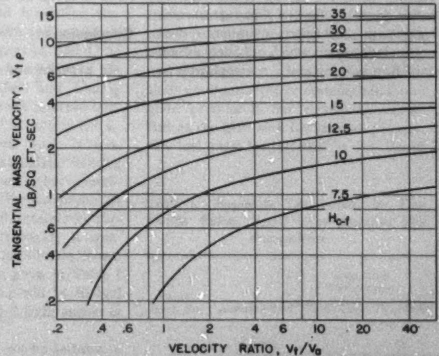


Fig. 13. Chart of average heat-transfer coefficients on pole face

The shape of the curves for consistent values of the tangential mass velocity, $V_t\rho$, is typical of the test data. These curves may represent conditions of constant speed of rotation and air density with varying air-flow rate, or conditions of varying speed and density with varying air-flow rate, but with the product of speed and density remaining constant. The lines for constant velocity ratios shown in Fig. 11 are parallel. The maxima occur at $(V_t/V_a)=13$, the minima at 6.3, and the inflection points at 7.5. The significance of the last value is its association with a transition in the flow pattern. With increasing values, the predominance of the rotational velocity increases. At decreasing values, the predominance of the axial flow velocity increases. When $(V_t/V_a)>13$, the air-flow rate becomes of minor importance. As indicated in Fig. 10, only when $(V_t/V_a)<3$ does the rotational effect become relatively insignificant. These observations are substantiated by the distributions of local heat-transfer coefficients (of the same type shown in Fig. 8) for the entire range of the test data.[1]

POLE FACE SURFACE

The average heat-transfer coefficients of the pole face are correlated with the tangential mass velocity as shown in Fig. 12. For the range of the test data where $0.2<V_t<15$ and $0.2<(V_t/V_a)<50$, the correlation may be expressed by

$$H_{e-f}=5.86(V_t\rho)^{0.65}+5.2(V_t/V_a)^{-0.376} \qquad (7)$$

The presentation of equation 7 in Fig. 13 permits direct determination of heat-transfer coefficients. The shapes of the curves indicate that the heat transfer is principally a function of the tangential mass velocity. Only at small velocity ratios does the effect of flow rate become significant. Therefore, equation 7 does not include a term which only depends on axial flow velocity and is not applicable to standstill conditions. Because of the predominance of the tangential component, standstill and rotational data do not show a pronounced relationship which would permit their correlation.

APPLICATION OF HEAT-TRANSFER DATA

The correlation plots of Figs. 11 and 13 are specific for the test configuration, which represents an idealization of an actual machine rotor. The tangential mass velocity, which is proportional to the speed of rotation and the air pressure, promotes air circulation in the open V-shaped passages between the poles of the test configuration and substantially increases heat transfer at higher air density. Therefore, the determination of heat-transfer coefficients for rotors with completely or partially closed passages by means of Figs. 11 and 13 may yield adequate estimates only for high-altitude (low-pressure) conditions.

For air passages of the same type as those of the test rotor, the data may be applied in the ranges of V_t and (V_t/V_a) covered in the measurements. The passage cross sections would have to be triangular or symmetrically trapezoidal and their equivalent diameter (four times the cross-sectional area divided by the cross-sectional perimeter) and their ratio of the passage base width at the rotor periphery to the passage radial height should be no more than 50% greater or 35% smaller than those of the test configuration, in order to lead to predictions of heat-transfer coefficients within ±15%. The range of the equivalent diameter would be 0.45 to 1 inch, and of the base width–radial height ratio, 0.7 to 1.7. Other tests made in the same study[1] with pole aspect ratios (length to width) in the range from 7.2 to 12.8 indicated no significant effects on heat transfer.

Conclusions

An experimental method is presented which provides means for the determination of heat-transfer coefficients on models of blast-cooled rotating machine components and can, therefore, be utilized as a tool for the development of thermally improved component and air passage configurations.

The data obtained on one specific model of a salient-pole alternator field structure have indicated that substantial variations of local heat-transfer coefficients may occur on pole surfaces. The positions of minimum heat-transfer coefficients indicate potential hot spots in insulations of appreciable thickness. The effects of air density and speed of rotation on heat transfer from pole side surfaces and faces are great, particularly when the flow passages are open at the rotor periphery. The air-flow rate and the ratio of tangential surface velocity to air approach velocity characterize the heat transfer from the side surfaces. The tangential surface velocity and the air density mainly determine the heat transfer from the pole faces.

References

1. SURFACE HEAT TRANSFER COEFFICIENTS OF SALIENT POLES IN A BLAST-COOLED ALTERNATOR, F. S. Tse, W. Robinson. *WADC Technical Note 56-123, ASTIA Document No. 110418*, Armed Services Technical Information Agency, Wright Air Development Center, Dayton, Ohio, Apr. 1956.

2. INVESTIGATION OF VARIATION OF POINT HEAT UNIT TRANSFER COEFFICIENTS AROUND A CYLINDER NORMAL TO AN AIR STREAM, W. H. Giedt. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 71, 1949, pp. 375–81.

3. INVESTIGATION OF POINT UNIT HEAT TRANSFER COEFFICIENTS FOR LAMINAR FLOW OVER AN INCLINED FLAT PLATE, R. M. Drake. *Journal of Applied Mechanics*, New York, N. Y., vol. 16, 1949, pp. 1–8.

4. LOCAL HEAT TRANSFER COEFFICIENTS ON THE SURFACE OF AN ELLIPTICAL CYLINDER IN A HIGH-SPEED AIR STREAM, R. A. Seban, R. M. Drake. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 75, 1953, pp 235–39.

5. FREE-CONVECTION HEAT TRANSFER FROM A ROTATING HORIZONTAL CYLINDER TO AMBIENT AIR WITH INTERFEROMETRIC STUDY OF FLOW, G. A. Etemad. *Ibid.*, vol. 77, 1955, pp. 1283–89.

Thermal Consideration of Generators in High-Speed Aircraft

ALEXANDER KUSKO
MEMBER AIEE

THE PROVISION of electric power in aircraft becomes increasingly difficult as the speed and altitude of flight increases. The relatively large values of electric power generated per pound of equipment in subsonic aircraft are achieved by employing the low-temperature environmental air, directed as a blast through the generator, to remove the generator heat losses and limit the internal temperature. As actual and projected flight speeds increase, the air blast becomes less effective as a cooling medium, because the stagnation temperature of the air approaches and exceeds the allowable temperatures of the insulating materials, bearings, and brushes of the generator. To handle the cooling problem, the present trend in generator design is to introduce new materials which are able to operate at higher temperatures, to employ a circulating coolant such as oil, and an external heat exchanger, and to depend upon the evaporation of liquids for cooling.

Under the sponsorship of the Air Force a group in the Energy Conversion Laboratory at the Massachusetts Institute of Technology has been studying the thermal problems of electric generating systems for aircraft operating at speeds up to Mach 4, four times the speed of sound, or approximately 2,700 miles per hour, and altitudes up to 100,000 feet. The group has studied two phases of the problem; one phase is the determination of the maximum temperature at which the electromagnetic rotating machine method of energy conversion is feasible. The second phase is the determination of optimum combinations of generators and cooling systems that employ presently available materials and components.

Problems Introduced by High-Speed Flight

The basic thermal problem of electric generators in high-speed aircraft, based upon extending the techniques presently used at lower speed and altitudes, is the rapid rise with speed of the stagnation temperature of air taken on board as a heat sink for generator losses. The stagnation temperature of cooling air at four flight design points is shown in Table I; note that the temperature rises from a value of 27.7 C (degrees centigrade) at Mach 1 and 30,000 feet altitude, to 724 C at Mach 4 and 100,000 feet altitude. Obviously, the materials employed in aircraft generators must be capable of operating at temperatures above the stagnation temperature of blast air if the air is to serve as a heat sink.

By comparison with the stagnation temperatures of blast air over this flight range, the maximum temperatures at which typical materials and components can be operated in aircraft generators today are shown in Fig. 1. It is interesting to note that insulating materials which are available and capable of operating at temperatures of 250 C no longer set the temperature limits of the machine; silicon rectifiers and bearing lubricants are now limiting internal temperatures to the range of 125 to 150 C. Of course, development work is continuing on dielectric materials, bearing lubricants, rectifiers, and other components, so that in time the maximum internal temperatures will certainly be permitted to increase. However, a comparison of the temperatures limits shown in Fig. 1 with the data given in Table I shows why blast cooling will be limited to the Mach-2-to-Mach-3 flight region in the near future.

Should blast cooling not be suitable, a number of alternate methods appear for transferring the generator losses to a heat sink. However, these proposed methods may exact a large penalty from the aircraft performance just to provide generator cooling. In Fig. 2 a set of curves is shown relating the percentile decrease in maximum range of a typical aircraft to the sum of the dead weight plus the weight equivalent of the drag imposed by the cooling system. The optimum system for a flight design point or flight profile, is the combination of generator and cooling system that imposes a minimum over-all penalty on the aircraft performance, as measured by a comprehensive penalty-factor method.

Generating System Development

In the face of rising stagnation temperatures of blast air, several approaches may be followed to achieve generating systems for high-speed aircraft. The development of higher-temperature materials and components will result in a gradual increase with time of the maximum speed-altitude design point at which blast cooling can be used. The higher machine temperatures will also reduce the temperature differential through which an external cooling system must carry generator losses to an arbitrary heat sink; i.e., an increase in internal machine temperatures will simplify and reduce the weight of external cooling systems for generators.

A second approach to the development of generating systems for high-speed aircraft is the introduction of so-called environment-free machines. These machines utilize an evaporating liquid as a heat sink, with the liquid either directly inside the machine or in an associated water boiler. The weight of these cooling systems is proportional to flight time; furthermore, they introduce certain other problems that limit their application as a long-range solution to the power generation problem. An essentially environment-free system that shows great promise is one employing engine fuel as the heat sink for generator losses; the fuel may be retained in the liquid

Table I. Stagnation Temperatures for Four Design Points, Based Upon Standard Hot Atmosphere

Speed, Mach No.	Altitude, Feet	Stagnation Temperature	
		Centigrade	Fahrenheit
1	30,000	27.7	82
2	50,000	146	295
3	70,000	380	716
4	100,000	724	1335

Paper 57-463, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted December 26, 1956; made available for printing March 14, 1957.

ALEXANDER KUSKO is with the Massachusetts Institute of Technology, Cambridge, Mass.

The research reported in this paper was made possible through the support extended the Massachusetts Institute of Technology, Servomechanisms Laboratory, by the United States Air Force (Equipment Laboratory, Wright Air Development Center) under Contract No. *AF 33(616) 2921*. M. I. T. Project Number *DSR-7352*. The assistance of Dr. E. O. A. Naumann is gratefully acknowledged. The work was carried out by the following group: P. Roy Choudhury, P. Hjertberg, R. M. Moroney, T. H. Putman, and A. K. Sen, all of the Massachusetts Institute of Technology.

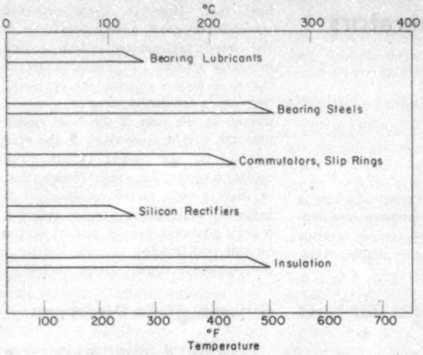


Fig. 1. **Temperature limits of presently available materials and components for generator construction**

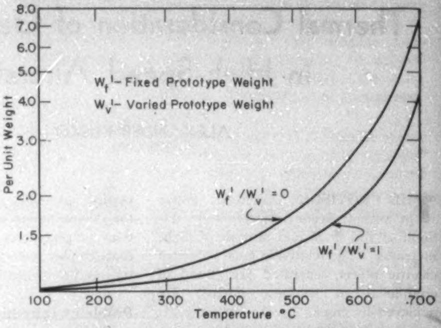


Fig. 3. **Per-unit weight of aircraft generator versus temperature for Curie temperature of 740 C. W_v'/W_f' = ratio of those weights which must change with temperature (i.e., iron) to those which remain fixed (bearings, etc.)**

state in passing through the heat exchanger, or permitted to change to the vapor state.

A third approach to generator cooling is the use of ram air, which is obtained as either blast air, or engine-compressor bleed air, as the heat sink for generator losses. The ram air is utilized for direct heat transfer by processing it with expansion turbines and heat exchangers to lower its temperature below the maximum allowable temperature of the generator, or by employing a refrigeration system for transferring generator losses up to the temperature level of the ram air. While the refrigeration system appears to be comparatively heavy, the ram-air processing systems are comparatively light and appear promising for high-speed aircraft. The ram-air processing system has the further advantage that generators designed for conventional blast-cooling operation can be inserted directly into these system without further development work.

Materials Considerations

Electric power generating equipment is synthesized from a combination of magnetic, dielectric, conducting, structural, and lubricating materials. Each of these materials suffers a downgrading in characteristics with increased temperature; some are more critical at this state of the art than others. Magnetic materials undergo a decrease in magnetization with increase in temperature, which culminates in a loss of ferromagnetic property at the Curie temperature. Typical Curie temperatures are approximately 750 C for iron and 1,100 C for cobalt.

The deterioration of dielectric materials with temperature is well known, and results in loss of both mechanical and dielectric strength. The presently available pliable dielectric materials are limited in temperature to about 260 C for the silicons, although Teflon can be operated above 300 C under certain conditions. To reach insulation operating

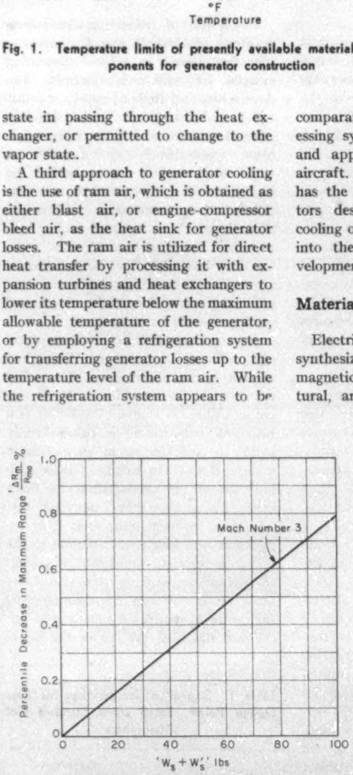


Fig. 2. **Aircraft range penalty due to addition of cooling system. Per-cent decrease in maximum range plotted against sum of cooling system dead weight W_s and dead-weight equivalents of drags imposed by cooling system W_s'. Based on empty aircraft weight of 10,000 pounds**

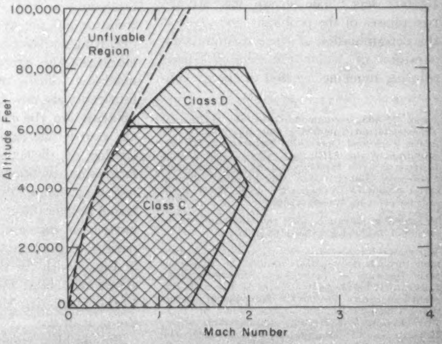


Fig. 4. **Operating region of blast-cooled class C and class D generators**

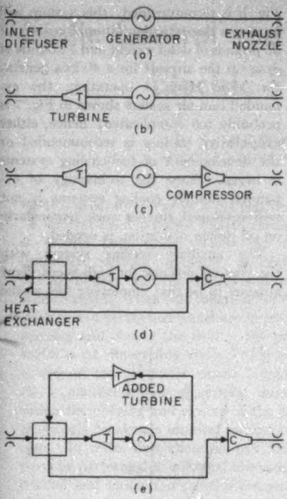


Fig. 5 (left). Development sequence of expanded ram-air topology

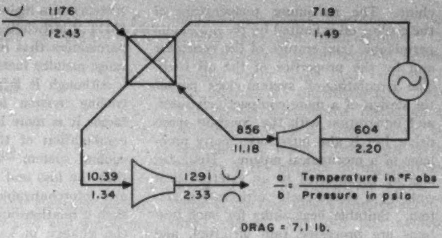


Fig. 6. Optimum Mach 3 system parameters

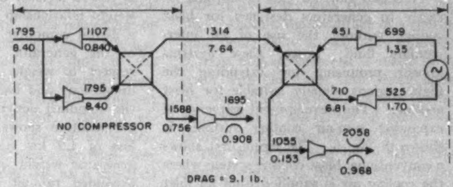


Fig. 7. Optimum Mach 4 system parameters

temperatures of 1,000 C and over, ceramic-type materials must be employed at present. However, these materials are solid, not pliable, and may require a new philosophy of machine design and construction for their use.

Two problems are encountered with conducting materials for high-temperature operation; first, the increase in resistance with temperature causes an increase in I^2R loss and a more difficult cooling problem. For example, copper has twice the resistance at 235 C that it has at zero C. The second problem is a tendency for surface oxidation of the copper to produce additional high resistance; techniques such as nickel plating the conductor have alleviated this problem.

The temperature limitation of structural materials does not appear so markedly in the range of temperatures encountered for generator operation. The structural materials probably cause more difficulty over the required wide temperature range because of bearing fit and thermal shock, rather than because of strength at a steady-state temperature. Lubricating materials are a present limitation in the maximum permissible temperature of generators; new types of lubricants and bearing systems are being developed to raise this temperature limit.

Should improvements on insulating materials and lubricants permit generator temperatures to be raised from the point of view of design, then the temperature effect on the magnetic materials (Curie point) will start affecting the size and rating of high-temperature rotating generators. The decline in the saturation flux density of the ferromagnetic material as its temperature approaches the Curie point requires that the magnetic-circuit cross section be increased to obtain a given generator rating with temperature. An example of how generator weight can be expected to increase with temperature, because of the magnetic-circuit requirements is shown in Fig. 3. Obviously, the development of ferromagnetic materials with higher Curie temperatures is necessary to conserve the weight of high-temperature machines.

Since the improvement of materials and techniques has permitted the maximum operating temperatures of aircraft generators to increase with time at a slow and expensive rate, it is interesting to note the effect of this increase on the flight regions of generator application to aircraft. In Fig. 4 are shown the steady-state operating regions in the altitude-speed plane of blast-cooled class *C* and class *D* generators; the class *C* generator is designed to operate with a maximum inlet air temperature of 120 C and the class *D* at 250 C. The class *D* generator is not yet available from industry. Whereas the class *D* generator will represent a major design step in maximum operating temperature, the resulting enlargement of the steady-state operating region of aircraft equipped with this blast-cooled generator is very moderate. For example, the blast-cooled class *C* generator can be operated in aircraft up to a speed of Mach 2 at 40,000 feet, with a maximum altitude limitation of 60,000 feet; the class *D* generator, on the other hand, can be operated in aircraft up to a speed of Mach 2.5 at an altitude of 50,000 feet, with a maximum altitude of 80,000 feet. However, the use of a higher-temperature generator will reduce the weight and complexity of an external cooling system, should one be necessary, and will also reduce the drag imposed on the aircraft when the generator is blast cooled.

External Cooling Systems

The alternative approach to developing high-temperature direct blast-air-cooled generators is to utilize generators of presently available materials and to design for them external cooling systems for transporting their electrical losses to a preselected heat sink. The present trend in generator design, where stagnation temperatures and other factors do not permit direct blast cooling, is to employ a circulating-oil system in which the oil carries the heat from the generator to an external heat exchanger. This approach does not solve the fundamental cooling problem, but transfers the problem from within the machine to outside the machine. The maximum temperature of the cooling oil is limited by the maximum permissible temperature of the generator and by the properties of the oil itself. The circulating-oil system does permit the design of a more compact generator, and integration with the constant-speed drive, but it also introduces many problems of a mechanical nature. Heat can also be transported to an external heat sink by a closed-cycle refrigeration system. Suitable heat sinks for such systems are processed ram air, fuel, and evaporants.

Processed ram air can be applied directly to generators designed for blast cooling without the use of an intermediate fluid; hence, these systems appear promising for extending the flight regions already suitable for blast cooling.[1] The development sequence of expanded ram-air cooling systems is shown in Fig. 5. The figure first shows a conventional blast-cooling system where the cooling air is supplied from an inlet diffuser and is exhausted through a nozzle; the next step shows the addition of a turbine for expanding the ram air to obtain a decrease in temperature and in pressure of the air. The third step in the evolution of the cooling system is the addition of a compressor on the exhaust side of the generator; to load the expansion turbine mechanically and increase the thrust from the exhaust air. Since the exhaust air is limited to the maximum internal temperature of the generator, under certain flight conditions the exhaust air may be limited to a temperature below that of the inlet ram air. In that case, as shown in Fig. 5(d), a regenerative air-to-air heat exchanger may be profitably added to the system. The final cooling system topology shows the addition of a second turbine on the same mechanical shaft as the first two turbines so as better to utilize the generator exhaust air.

These expanded ram-air systems appear so promising for generator cooling, particularly at the Mach 3 design point, that they were studied in detail. Optimum design conditions were found. The system shown in Fig. 6 is an optimum expanded ram-air system for Mach 3 operation at 70,000 feet altitude of a single 40-kva generator. A more complicated system for Mach 4 and 100,000 feet altitude operation with a 40-kva generator is shown in Fig. 7. The optimum parameters were obtained by programming the design of these systems on a high-speed digital computer (MIT Whirlwind) and selecting the set of parameters that resulted in the smallest range-penalty factor.

Although it is possible to optimize a cooling system for minimum penalty factor it is more logical to optimize the combination of the generator and the cooling system. It is well understood that the loss and weight of a generator are interchangable in a design sense; such a relationship is shown in Fig. 8 as the change in loss and the change in weight from standard conditions of a class B 40-kva aircraft alternator. This loss-weight relationship was determined with a digital computer on which the design of the generator was optimized with respect to weight for perturbations in dimensions and losses above and below the standard design. For example, the computation shows that an increase of loss of 1.2 kw results in a saving of generator weight of 15 pounds. Since expanded ram-air cooling systems appear to accommodate an increase in power loss capacity without a proportional increase in weight, an optimum system may consist of a generator having greater loss but lighter in weight than the existing blast-cooled machine and the appropriate expanded ram-air cooling system.

Conclusions

By employing ram-air processing equipment, it appears that existing aircraft generators can be operated over the range of flight speeds from Mach 1 to Mach 4 and appropriate altitudes, within acceptable penalty factors to the aircraft. The raising of maximum temperatures of generators by using new materials and components will reduce penalty factors to aircraft but will not markedly extend the flight regions in the speed-altitude plane for blast cooling. Extensions of the flight region will certainly require some type of auxiliary cooling system for the generator or for the group of components of which the generator is a part.

The recommended method for cooling generators for flight speeds from Mach 1 to Mach 2 is blast cooling, using either the presently available class *C* type, or the anticipated class *D* type, which is to operate with an inlet air temperature of 250 C. For Mach 3 aircraft operation the expanded ram air system shown in Fig. 6 is recommended: this system will impose a theoretical additional penalty of 35 pounds of dead weight and 7 pounds of drag on the aircraft for a 40-kva generator. For Mach 4 operation, the expanded ram-air system shown in Fig. 7 is probably too complicated; hence, either evaporative cooling is recommended or the development of fuel-cooling systems is urged. Since the technology of expanded ram-air cooling systems is not well developed, further work, particularly on off-design operation, is needed.

The auxiliary cooling system with the greatest promise for high-speed flight appears to be that employing fuel as the ultimate heat sink. The system has limited capacity if the fuel remains in the liquid phase; however, the system is very promising if the fuel is allowed to vaporize in the heat exchanger so that the latent heat of vaporization is utilized. Finally, it appears that the ultimate solution to the problem of generating electric power in high-speed high-altitude vehicles is the development of new energy-conversion processes that can better employ existing materials in the high-temperature environment encountered.

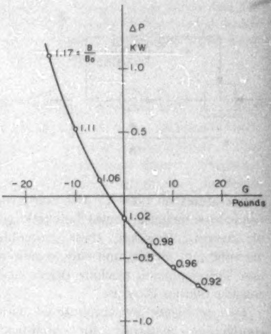


Fig. 8. Change in loss as function of change in weight for aircraft generator

Reference

1. RAM-AIR COOLING SYSTEMS FOR AIRCRAFT GENERATORS, Richard M. Moroney. *AIEE Transactions*, vol. 76, pt. II, Sept. 1957, pp. 217–21.

Static Control and Protection of Aircraft Electric Power Systems

R. E. HULSEY
ASSOCIATE MEMBER AIEE

N. F. SCHUH
ASSOCIATE MEMBER AIEE

THE generation end of a conventional aircraft electric-power system consists of several basic units. These are the generator, voltage regulator, generator circuit breaker, and the control and protection panel. The basic functions of the control and protection panel are those of opening and closing the field of the generator, and supervising the opening and closing of the generator circuit breaker. In a parallel-generator system, supervising the opening and closing of the bus tie circuit breaker is another basic function. In order that these control functions be performed in a logical manner, a control and protection panel must sense voltages and currents throughout the electric system. These various voltages and currents are measured, compared with standards, time-delayed if necessary, mixed with each other in logic circuits, amplified, and finally used to perform the necessary control functions. The panel protective functions include such circuits as overvoltage protection, undervoltage protection, overexcitation protection, underexcitation protection, underspeed protection, stability protection, and differential current protection. The panel control functions include lockout circuits to prevent cycling of circuit breakers, and a circuit for automatic paralleling of generators.

A conventional control and protection panel is an assembly of electromechanical relays, resistors, transformers, and rectifier stacks. Because closely calibrated electromechanical relays are used, this assembly is usually suspended on vibration isolators, to overcome the problems caused by mechanical environment.

A number of problems are encountered with conventional equipment such as described. The major problems result from the necessity of using calibrated relays. Voltages and currents around the system are measured and sometimes rectified, but the mechanical relay itself in each protective or control circuit is the point of calibration, and thus the critical point of the circuit. Each relay must be set so that its operating point is fixed at the desired value. Not only is the pickup value of the relay always important, but in some cases the drop-out point is also important. This is especially true in the case of an undervoltage protection relay. In this case the relay must drop out at a specified value of undervoltage, e.g., at 100 volts line-to-neutral on a nominal 115-volt line-to-neutral system. It is obvious that the pickup value of the relay must be at some point below the nominal operating voltage of the system. To take into account variations in setting of the voltage regulator, the pickup would probably be set below 110 volts. This requires a closely calibrated relay which has a pickup and a dropout between 100 and 110 volts.

In some cases, alternating voltages or currents are measured and used to operate a-c relays directly. These a-c relays have the advantage of not requiring temperature compensation, but have the disadvantage of requiring many more volt-amperes for operation than a d-c relay of an equivalent contact rating. These calibrated relays must be set not only to very accurate pickup and sometimes drop-out points, but these operating points must be held under the various environmental conditions specified for aircraft application. One of the most troublesome of these conditions is temperature, because the resistance of copper coils used on relays changes considerably with temperature, and therefore must be compensated for. The mechanical, environmental conditions are also troublesome, because calibrated relays are especially sensitive to vibration, shock, and acceleration. The use of vibration isolators makes it possible to use these relays on aircraft, but with a sacrifice in operating characteristics, and the constant threat of nuisance operation.

On conventional control and protective panels, thermal relays are used to obtain required time delays; however, they have several important disadvantages. One is the amount of power required, which is usually many times that required for an equivalent electromagnetic relay. This power requirement places an excessive load on the d-c control power bus, and can also leave a fixed load on the battery of the airplane during shutdown, unless special provisions are made. In general, the time delays required are to be fixed delays. The d-c bus available generally has a wide variation of voltage, which means that the thermal relay time delays vary with this voltage supply. Another disadvantage of a thermal time delay relay is that after operation it must cool off before it is ready to operate with the same time delay again. An attempted operation of the thermal relay before it has cooled off results in a proportionately shorter time delay.

The power required to be taken from sensing circuits in the conventional control and protection panel is very often a disadvantage. The a-c and d-c relays which must be operated generally require enough power to cause a definite decrease in the possible sensitivity of the sensing circuit used. This, of course, results in power wasted, and heat generated which must be dissipated in some manner. The dissipation of heat is becoming a more important problem in each new aircraft application. Fault protection, which makes use of various types of sequence filters, is coming more into use in new aircraft applications. If enough power to operate an electromechanical relay is required from these filters, the components of the filters become very large, and the sensitivity lessens. This type of sensing is increasingly important, and therefore a means of using small components in the sequence filters and of still maintaining high sensitivity is becoming equally important. Current transformer loops generally work into a-c relays which require 3 or 4 volt-amperes for operation. This means that the current transformers must be able to handle this load.

Temperature compensation is a problem in any type of control device and becomes a more difficult one with the power requirements of the sensing circuits. The more current or voltage which must be handled, the larger must be the temperature compensating components.

New Approach

The limitations of present methods of control and protection indicate that a new approach should be made; that

---

Paper 57-469, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 8, 1957; made available for printing April 11, 1957.

R. E. Hulsey and N. F. Schuh are with the Westinghouse Electric Corporation, Lima, Ohio.

The authors wish to acknowledge the work performed by J. R. Reeder in the development and design of the equipment described herein.

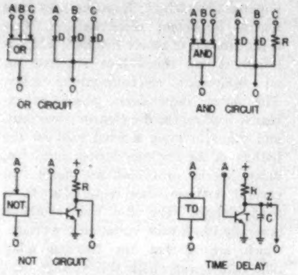


Fig. 1. Basic logic circuits and time delay circuits

See Table I for Legend

is, the indications are that a means should be found to perform the desired functions without those limitations discussed. Several new approaches to the problem have been studied. One method has been selected, and a control and protection panel has been designed which makes use of this new approach. The results are described in this paper.

For the study of a new approach to a problem, the availability of new materials should be studied. Very often a new approach which appears desirable is not possible for lack of new materials to accomplish the job. In this specific case the many new devices, available as a result of developments in the semiconductor field, were studied, and it was decided to make a new approach to the problem of control and protection circuitry by making use of such new semiconductor devices.

Rectification of alternating current to direct current has been used on panels since the first a-c system built. In most cases, however, rectification has been performed by selenium rectifiers. Semiconductor rectifiers have made possible not only a significant reduction in size, but an increase in reliability, life, and maximum operating temperature. Silicon rectifiers, for example, do not exhibit an aging characteristic, and, because of their smaller size and sealed construction, they are not as sensitive to environment. The availability of semiconductor rectifiers of various sizes and shapes has made possible the free use of these devices in new circuits without a significant penalty in size and weight. Silicon rectifiers or diodes are used: for straight rectification, replacing common selenium rectifier stacks; as blocking diodes in logic circuits; and as zener diodes, to detect when a predetermined voltage level is reached. Silicon transistors, another of the semiconductor devices now available in quantity, although still limited in power-handling capabilities, are used both as amplifiers and as mixing or logic devices. There are several semiconductor devices exhibiting a negative resistance characteristic which are in the developmental stage and which will undoubtedly turn out to be very useful in control and protection circuitry. This negative resistance characteristic makes possible their use in bistable or triggering circuits.

With any discussion of new materials must be mentioned the progress made in the miniaturization of standard circuit components. Very small, high-quality wire-wound resistors, with different temperature coefficients of resistance, may be obtained. Miniature potentiometers, or voltage dividers, and variable resistors, allowing adjustment without sacrificing very much in size and weight are now obtainable. The use of tantalum has made available high-quality capacitors with large values of capacitance in very small packages.

Table I. Legend

AMP	Amplifier
AP	Automatic paralleling
BV	Bus voltage
CL	Close
DC	Direct current
DP	Differential protection
EC	Exciter ceiling
GC	Generator control
LAD	Latched analog-to-digital converter
LO	Lockout
OE	Overexcitation
OV	Overvoltage
SP	Stability protection
TD	Time delay
TR	Trip
UE	Underexcitation
US	Underspeed
UV	Undervoltage

Another item which must be mentioned in any discussion of new methods or materials is the rapidly increasing use of printed circuitry. Printed circuits now hold an important position in the field of electronics or electric control. Aside from their obvious advantages, printed circuits provide a means of mounting the many small parts such as silicon diodes, transistors and miniature resistors, and eliminate the possibility of wiring errors. The availability of new materials and devices, the progress which has been made in the miniaturization of parts, and the development of printed circuits, have led to this new approach to circuit design in control and protection for aircraft electric-power systems.

Circuit Design

The circuit design philosophy which has been developed on the basis of the comments in the preceding paragraphs is that all functions of the control and protection panel should be performed statically if possible, i.e., with no moving parts; and further, that as many functions as possible should be performed at a very low power level. This means that the functions of sensing, logic or mixing, time delays, and lockouts should all be accomplished at a very low power level. The power level is increased, then, only after all the logic functions are performed, and when it is finally necessary to perform a control function. At this stage amplification is accomplished, so that the control functions, such as operation of circuit breakers (because they require larger amounts of power) can be performed.

The circuitry of a control and protection panel may be broken down into several functional groups and discussed separately. Sensing circuits are those circuits which measure the various voltages and currents throughout the electric power system and convert them into signals which can be used by the subsequent circuits in the panel. In general these sensing circuits are merely transformers and rectifiers to provide d-c signals which are proportional to alternating voltages and currents. As required, some sensing circuits may include frequency-sensitive filters or some type of sequence filter. All sensing is accomplished at a very low power level, making possible the use of very small parts.

Logic circuit is the name given to the various computer-type data-processing circuits which perform the mixing functions required in the panel. These

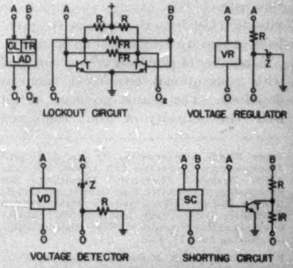


Fig. 2. Circuits for basic control functions

See Table I for Legend

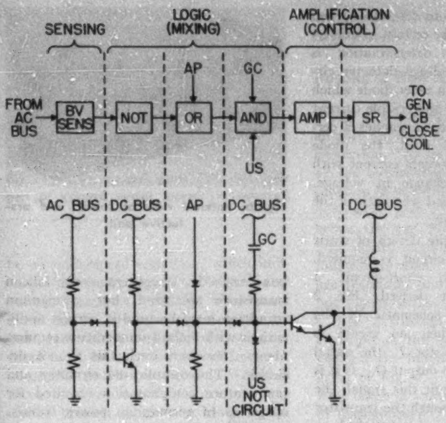

SENSING
LOGIC (MIXING)
AMPLIFICATION (CONTROL)
FROM AC BUS
BV SENS
NOT
AP
OR
GC
AND
US
AMP
SR
TO GEN CB CLOSE COIL
AC BUS
DC BUS
AP
DC BUS
GC
DC BUS
US NOT CIRCUIT

Fig. 3. Development of control function using basic logic and control circuits

See Table I for Legend

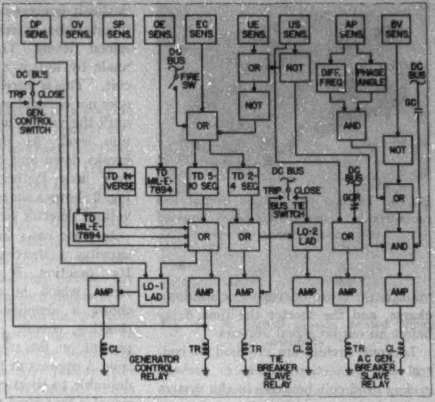

DP SENS.
OV SENS.
SP SENS.
OE SENS.
EC SENS.
UE SENS.
US SENS.
AP SENS.
BV SENS.
DC BUS
TRIP
CLOSE
GEN. CONTROL SWITCH
DC BUS
FIRE SW
OR
NOT
NOT
DIFF. FREQ
PHASE ANGLE
DC BUS
GC
AND
NOT
OR
TD INVERSE
TD MIL-E-7894
TD 5-10 SEC
TD 2-4 SEC
DC BUS
TRIP
CLOSE
BUS TIE SWITCH
DC BUS
GCR
OR
TD MIL-E-7894
OR
OR
LO-2 LAD
OR
AND
AMP
LO-1 LAD
AMP
AMP
AMP
AMP
AMP
GL
TR
TR
CL
TR
CL
GENERATOR CONTROL RELAY
TIE BREAKER SLAVE RELAY
AC GEN. BREAKER SLAVE RELAY

Fig. 4. Function diagram of typical a-c control and protective unit for parallel system

See Table I for Legend

logic circuits include AND, OR, and NOT circuits. Fig. 1 shows a simplified schematic diagram of each of these types of circuits. An OR circuit is one which requires that only one of its inputs be present before an output is obtained; that is, an output will appear at 0 when a signal is present at input *A*, *B*, or *C*. An OR circuit may be used to provide the signal for opening the field of the generator. There may be three conditions, any one of which might trip the field of the generator open. These are an overvoltage condition, a ground fault condition, or an undervoltage condition. The outputs of the three sensing circuits which measure these three conditions would be fed into the OR circuit, and if any one of these conditions were present the OR circuit would have an output to open the field. An AND circuit is one which requires that all of its inputs be present before an output signal is obtained; that is, a signal must be present at inputs *A*, *B*, and *C* before an output signal appears at 0. An AND circuit may be used to control the closing of a generator circuit breaker. Before this circuit breaker is closed, it might be required that three conditions be met; for example, correct frequency, correct voltage level, and the ground power plug disconnected. Signals from sensing circuits which measure these three conditions would be fed into an AND circuit, and no output would be obtained to close the breaker until all three conditions were met. A NOT circuit is one which has an output when there is no input, and does not have an output when there is an input; that is, an output signal will appear at 0 when no signal is present at *A* and vice versa. This type of circuit is used where a reversal of signal is required. A NOT circuit can be used to determine that there is not an underspeed condition. If there is not an underspeed condition, a signal would be available at the output of the NOT circuit.

A number of time delays are required in any control and protection panel, to provide proper co-ordination of control functions. These time delays may be divided into fixed time delays and variable time delays. Time delays are obtained in this design by making use of resistor-capacitor (*R-C*) circuits. The charging time of a capacitor, when connected to a voltage source through a resistor, provides a reliable time delay. Fig. 1 shows the simplified schematic of a fixed time delay. The time delay is determined by the value of the capacitor and the resistor, and the amplitude of the available supply voltage. Since the supply voltage available is generally not constant, a regulated voltage supply is used for these *R-C* time delay circuits. This is not difficult to accomplish, since the time delay circuits require very little power. Referring to Fig. 1, it may be seen that removal of the signal at the input *A* turns the transistor off. The capacitor is unshort-circuited, and begins to charge through the resistor. When the charge on the capacitor reaches a certain point, the zener diode *Z* breaks down and conducts, presenting an output signal at *O*. Thus, an output is obtained from the circuit after a certain fixed time delay. This type of time delay circuit requires no cooling-off period. It can be seen that as soon as the signal is returned to the base of transistor *T*, the capacitor *C* may discharge immediately. This circuit might be the time delay used in conjunction with the undervoltage sensing circuit on the panel. As long as normal voltage appears on the base of the transistor there is no output from the circuit. When an undervoltage condition exists, the transistor has no signal on its base, and the *R-C* circuit begins to function. After a fixed time delay a signal appears at the output of the time delay circuit to perform whatever control function is required. As soon as voltage returns to normal, the *R-C* circuit is discharged and ready for another cycle. Variable time delays are also obtained by the use of *R-C* circuits, but in this case, instead of making use of a fixed voltage source, the voltage being measured or sensed is also used as a supply voltage for the *R-C* circuit. In the case of overvoltage protection, an inverse time-voltage characteristic is required. This is obtained by providing a direct voltage, proportional to the alternating voltage, directly to the *R-C* circuit. The higher the degree of over-

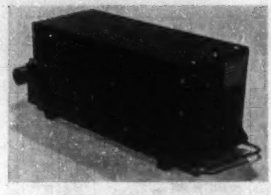


Fig. 5. External view of a-c control and protective unit as described in Fig. 4, mounted in quick-eject mounting rack

voltage, the faster the capacitor obtains a charge, and the shorter the time delay before an output signal appears.

Lockout circuits are required in control and protection panels to prevent cycling of circuit breakers in the system under fault conditions. In a conventional system, a lockout circuit may consist of several relays so interconnected that cycling is prevented. In static control, a type of bistable circuit is used. Fig. 2 shows a simplified schematic of a bistable circuit which is called a LAD circuit. This is an abbreviation for latched-analog-to-digital converter. It is a circuit in which one transistor or the other is conducting, but never both. The transistors are electrically latched, so that whichever one is turned on remains on, and holds the other one off until a signal is applied to the opposite transistor to turn it on. Normally the LAD circuit would be operating in a mode such that it had no effect on the normal operation of the panel. A fault on the electric system would apply a signal to the off transistor turning it on, short-circuiting out the resetting signal, and preventing automatic reclosure of the breakers.

Voltage regulator circuits are used to obtain constant direct voltages where only variable direct voltages are available. A simple voltage regulator circuit can be constructed, as shown in Fig. 2, with a zener diode and a resistor. The zener diode functions in a manner similar to the more familiar gas-filled voltage regulator tube. If, for example, the direct voltage available on the airplane varies between 20 and 29 volts, a simple regulator of this type, with the small load requirements of static circuitry, can be made to hold the variation of direct voltage to 18 to 20 volts. A regulated d-c bus, internal to the panel, is used to supply power to any of the static circuitry which may be voltage sensitive. All fixed time delays obtain their power from the regulated d-c bus.

Throughout the static circuitry of this design it is necessary to determine when direct voltages reach certain predetermined levels. This determination is made by use of a voltage detector circuit. This is merely a zener diode which does not conduct any appreciable current until the voltage across it reaches a certain level. At that level, the diode breaks down and conducts current with very little further change in voltage. Fig. 2 shows a simplified schematic of the voltage detector circuit.

In some cases in the design of static circuitry a shorting circuit is required. Its function is to short-circuit a signal which is not desired. Fig. 2 shows a simplified schematic of the shorting circuit. When no signal is present on the transistor *T*, the signal from A appears at the output *O*. If it is desirable to short-circuit this signal, the signal is by-passed through the transistor to ground by applying a signal to the base of the transistor. This type of circuit might be used to short-circuit a signal from a manually operated reset switch when it is desired that the system not be operated because of the existence of a fault.

After all the foregoing functions are accomplished, signals are available which indicate what final control functions are to be performed. These signals are at a very low power level, however, and must be amplified before they are of any use. Transistor amplifiers are used wherever possible in the static circuitry. At the present time static amplification is limited by the availability of transistors which can dissipate the power required. The final control device at present is a slave relay which is able to handle the operating requirements of the circuit breaker close and trip coils. A distinction must be made at this point between the slave relays used in this new approach and the calibrated relays which are being eliminated by this new approach. The limitations of the calibrated relays were covered in the first part of this paper. It should be understood that the slave relays mentioned here are very rugged relays whose coils are always either completely de-energized or fully energized. The relays are on-off devices and do not depend upon a calibrated operating point in any way. The transistor amplifiers are driven in such a way that the slave relay either has no signal applied to it or has more than enough signal applied for operation. The slave relays are standard ruggedized relays designed for aircraft application.

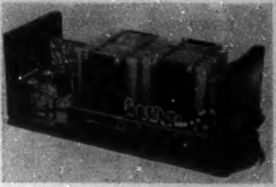


Fig. 6. Internal view of control and protective unit

The manner in which the transistors are used in this application should be described. In general, only silicon transistors are used, but germanium transistors may be used in certain applications with limited temperature requirements; however, even this is not desirable. The complicated circuitry and temperature compensation required for operation of germanium power transistors even at 71 C is undesirable. The general approach has been to use silicon transistors wherever possible, and when silicon power transistors are not available to use slave relays. It should be noted that in this application the transistors are always used as switching devices rather than as fixed gain amplifiers. No attempt has been made to depend on transistors having fixed gains or other characteristics, or even to temperature-compensate transistor stages so that their characteristics are fixed. Circuits have been designed so that under the worst conditions enough signal is present to drive the transistor stages; under all other conditions, the transistors are overdriven, but still have the same output signal level. Each transistor is either on or off as far as the subsequent circuit is concerned, whether the transistor is on because of having barely sufficient signal, or because it is being overdriven. Overdriving the transistors causes no trouble, since the current is limited by the proper selection of resistance values.

Typical Circuit

Fig. 3 shows the application of the design philosophy outlined in the foregoing to an actual circuit. This is the circuit which controls the closing of the generator circuit breaker in a typical aircraft electric-power system. In this case three conditions must be fulfilled before the circuit breaker is allowed to close. The generator control relay or field relay, *GC*, must be closed to indicate that the generator excitation circuit is closed, and therefore that the generator is producing voltage. There must not

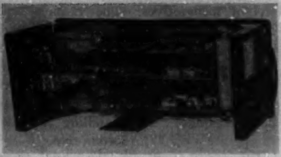


Fig. 7. Control and protective unit with part of printed circuit assemblies removed

be an underspeed condition, which indicates that the frequency of the generator voltage is correct. Finally, there must either not be any bus voltage on the tie bus, since this is a paralleled system; or there must be a proper signal from the automatic paralleling circuit. When all three of these conditions are met the circuit breaker is given a close signal. Fig. 3 indicates the block diagram of such a circuit with the simplified schematic of the same circuitry shown below. This figure shows how the various circuits described in the preceding paragraphs are connected together to perform one of the many functions of a control and protection panel.

Advantages of New Approach

Many advantages are realized by designing a control and protection panel which makes use of the design philosophy of the new approach outlined in this paper. The size of the final equipment is reduced. This is possible because of the decreased size of the individual components required. The reduction in size will be more apparent when power silicon transistors are available and thus allow the elimination of even the slave relays now used. The elimination of the need for vibration isolators further reduces the size of the final equipment in that no allowances need be made for movement of the chassis on its isolators. The use of semiconductor devices now available, and those which will be available in the future, allows extended application of this type of equipment in increasingly difficult environmental conditions. The power dissipation is decreased, an important point in new applications where the problem of dissipating heat becomes increasingly critical, and in evaluating the life of the components used.

Most important of all is the increase in the inherent reliability of such a piece of equipment. The greatest increase in reliability comes as a result of eliminating the use of calibrated electromechanical relays and the other components which are associated with these relays, such as large power-type adjustable resistors.

Results of New Approach

Fig. 4 is a block diagram of a specific aircraft application of static control and protection. As far as has been practical, all the blocks representing each type of function have been included in the same horizontal line. Sensing circuits are across the top with logic, time delays, logic again, amplifiers, and slave relays, following in order. Voltage detector, voltage regulator, and shorting circuits are not shown in this simplified diagram.

Figs. 5 through 8 illustrate the construction of the equipment designed to perform the functions of the block diagram of Fig. 4. The four photographs were taken in sequence at different stages of disassembly. Fig. 5 shows the completely packaged equipment, assembled in its standard quick-eject mounting rack. No vibration isolators are used, and all slave relays are hermetically sealed, with the result that no moving parts are exposed to the atmosphere. Fig. 6 shows the equipment with the cover removed. Most of the static circuitry is contained on the 13 small, printed circuit boards which can be seen standing on end. These are held in place by connectors on the lower end, and by through-bolts and spacers on the upper end.

Fig. 7 shows the panel with the through-bolts, spacers, and most of the printed circuit boards removed. The 13 connectors which hold the printed circuit boards are in turn mounted on, and soldered to, a master printed circuit board. This is a large board, with printed wiring on both sides, which is horizontal to the base. The slave relays and power diodes are mounted on the die-cast base underneath the master printed circuit board. The transformers are mounted at the connector end of the panel.

This view demonstrates the ease of maintenance which can be achieved by making use of printed circuits and miniature parts. The circuits on the small board are divided according to the function performed. For example, one board includes overvoltage sensing and time delay. Quick repair can be accomplished by removing the board with the faulty circuit and replacing it with a known good circuit. The plug-in feature also facilitates routine maintenance by allowing easy removal and checking of the various functional circuits.

Fig. 8 shows the arrangement of parts on a typical printed circuit board. The board illustrated contains the circuitry for bus voltage and automatic paralleling sensing. Parts mounted on the board include transistors, diodes, resistors, and capacitors. This picture illustrates the small size of the parts used in static circuit design. It can be seen that polarizing is accomplished by notches at the plug-in end of the board. Matching bars in the connector insure that boards are plugged into the correct connectors.

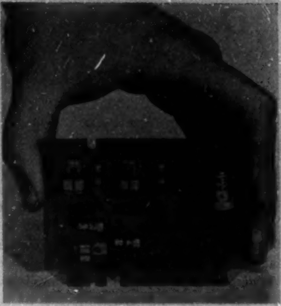


Fig. 8. A typical plug-in printed circuit assembly

Conclusions

The new design approach described in this paper produces an aircraft control and protective unit with the following improvements over conventional designs.

1. The elimination of closely calibrated relays by performing all sensing, logic, and mixing functions statically, and then amplifying, greatly increases the reliability of the control and protective unit.
2. The circuits require less power to perform their function, thereby allowing more sensitivity to be obtained.
3. By using the small rugged components, thus eliminating the necessity for vibration isolators, a reduction in size and weight is made.
4. The use of plug-in printed circuit assemblies allows ease of maintenance when necessary.

With these improvements, static control and protection brings one step closer the ultimate in aircraft electric-system performance.

References

1. A TRANSISTORIZED OVERVOLTAGE RELAY N. F. Schuh. *AIEE Transactions*, vol. 74, pt. II 1955 (Jan. 1956 section), pp. 407–10.

2. NEGATIVE-SEQUENCE VOLTAGE DETECTION ON AIRCRAFT A-C ELECTRIC SYSTEMS, R. D. Jessee, P. F. Boggess. *Ibid.*, vol. 75, pt. II, July 1956, pp. 171–75.

Design of Ignitron Tubes for Rectifier Locomotives

B. E. BELNAP
ASSOCIATE MEMBER AIEE

J. L. ZEHNER
MEMBER AIEE

H. E. ZUVERS
ASSOCIATE MEMBER AIEE

IGNITRONS are used to rectify a-c power to operate d-c drive motors on 22 rectifier locomotives. These rectifier locomotives represent the first large-scale application of the ignitron in railroad locomotive service. Ignitrons in railroad locomotive service are subjected to an electrical cycle which closely matches their inherent capabilities; although a motion environment of roll, pitch, vibration, and shock makes necessary certain mechanical design and construction techniques which are not required for ignitrons in other industrial applications.

Application

The electrical duty imposed on the ignitrons is defined by the locomotive drawbar pull required to meet train schedules. Locomotives in passenger service have frequent short-time torque peaks as the train accelerates. Fig. 1 shows a portion of a typical load cycle of a tube used on a rectifier locomotive in passenger service. The current and voltage requirements per tube are dependent on the number of tubes, the circuit selected, and the voltage level chosen, for a given locomotive load cycle. In general, the power-handling capabilities of the tube increase with voltage so the rectifier output voltage levels are as high as permitted by the traction motors. Ignitrons and other mercury pool devices have intrinsic capabilities which are as useful in locomotive service as in resistance welding. They have a capability based on average current carried, because the tube thermal losses are more closely proportional to average than to peak or rms current. The ratio of peak-to-average current can be very high for short periods.

In passenger locomotive service this short time capacity is used for starting and accelerating the train. A tube which is designed to handle the severe starting and accelerating duty can easily take care of the other portions of the load cycle, as has been proved by comprehensive tests. Tube electrical duty on freight locomotives closely approximates normal industrial rectifier ratings. Freight train accelerating times are long compared with the thermal time constants of the tube.

Electrical Design

Space is limited on locomotives, so the tube design must provide a maximum of power-handling ability in a given physical size. In any given size of vacuum and water jacket a number of grid-anode combinations are possible. Within practicable limits in 8-inch and 12-inch tubes there may be no grids, or 1, 2, or 3 grids. The highest power handling capability is provided by a 2-grid tube in the voltage range of a 2,000- to 6,000-volt peak. This voltage level was determined by the voltage limits of the traction motors.

Twelve 8-inch-diameter double-grid *GL-6504* ignitrons are used in each of ten 4,000-hp (horsepower) passenger locomotives, see Fig. 2. Twelve 12-inch-diameter double-grid *GL-6878* ignitrons are used in each of 12 3,300-hp freight locomotives. The 12-inch tube has roughly 180% of the power-handling capability of the 8-inch tube. It was chosen for the 3,300-hp freight locomotives because of their higher duty requirement. Twelve 12-inch *GL-6878* ignitrons have adequate power-handling capacity to operate a 6,000-hp locomotive.

Internal tube dimensions which effect ignitron characteristics were selected to give a maximum performance under all environmental conditions. The two graphite grids, which are shown in Fig. 3, with properly chosen thicknesses and grid hole diameter are partly responsible for the high-current and high-voltage handling ability of the locomotive ignitron over a wide range of temperature. An even more important consideration in getting a high tube capability is the coaxial cathode current connection. The locomotive ignitrons are the first commercially available ignitrons with a coaxial cathode current return which substantially reduces the effect of the tube's own current on the mercury arc in both the individual tube and adjacent tubes. There is no magnetic field outside a coaxial bus, therefore the arc in the locomotive tubes is comparatively unaffected by the current in the tube's own cathode return. The magnetic fields of the tubes due to the anode and cathode conductors are minimized by keeping the conductor above the tubes.

Vibration, shock, roll, and pitch affect electrical operation of the tube only if grid-anode spacings change or mercury from the cathode splashes onto the anode or grids. The mercury pool is restricted

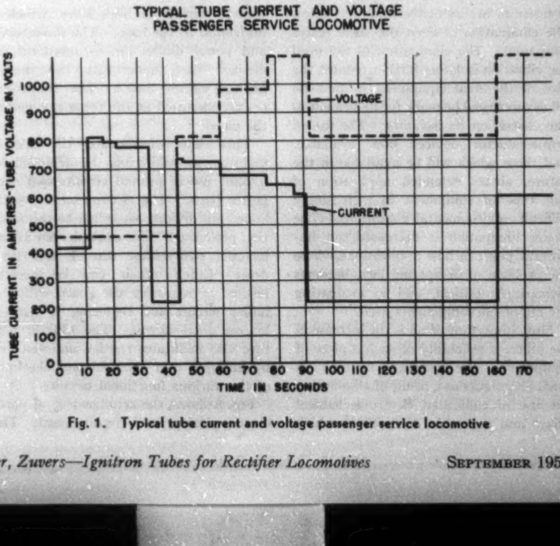


Fig. 1. Typical tube current and voltage passenger service locomotive

---

Paper 57-33, recommended by the AIEE Industrial Power Rectifiers Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N.Y., January 21-25, 1957. Manuscript submitted October 25, 1956; made available for printing February 6, 1957.

B. E. Belnap, J. L. Zehner, and H. E. Zuvers are with the General Electric Company, Schenectady, N.Y.

in motion so that there is no mercury splash. The tube has been tested and operates under shock and vibration conditions more severe than any encountered in railroad service.

Electrical tests on the tubes used in locomotives have been made on both single-phase and 3-phase equipment at 60 cycles but only on single-phase at 25 cycles. Three-phase 60-cycle operation is most severe and, subject to thermal maximums, it was found that a tube rated on a 3-phase basis could be rated about 20% higher when used in single-phase circuits at the same frequency. The permissible change in rating of the locomotive tubes due to the use of 25 rather than 60 cycles cannot be evaluated because a thermal limitation is reached at 60 cycles: see Fig. 4. The commutating capacity of the double-grid construction extends the possible rating beyond the thermal limit.

Mechanical Design, General

The general type of construction, shown in Fig. 5, and the materials used in the 12-inch *GL-6878* and 8-inch *GL-6504* locomotive tubes follow well-established ignitron practice which has been proved in thousands of industrial applications over a period of 18 years. Such features as the use of stainless steel in contact with the cooling water, the best grades of electronic graphite for anode, grids, and baffle, and the use of fernico and glass with edge-type seals, are considered standard.

Special Features for Locomotive Service

It was found to be both desirable and necessary to build in some special construction features because of the operating conditions peculiar to locomotive rectifiers, specifically, the continual vibration, shock, and swaying to which the tubes will be subjected. For instance, a special alloy is used for the bolts which clamp the grid-support insulator assemblies together because ordinary steel bolts would expand so much more than the ceramic insulators that they would allow some movement between parts when they are vibrated at normal operating temperature. While this movement in itself would not be great enough to change the grid spacings appreciably, it would allow a grinding action between the ceramic and steel parts which could shorten the useful life of the tube.

Whenever threaded parts are used in assembly of the tubes, the nuts are tack-welded or a positive locking device is used to prevent loosening from vibration.

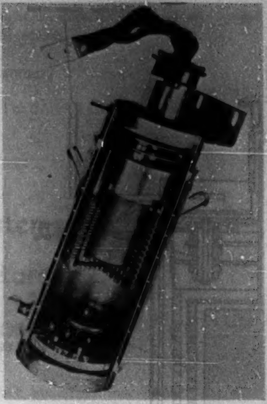


Fig. 2. Cutaway of GL-6504 Ignitron

Fig. 3. Graphite parts for GL-6504 and GL-6878

Every component and construction technique was pretested before the final design was selected and many departures from standard practice were found necessary to ensure long life in spite of vibration and shock.

The primary mechanical problem in designing the cathode was to provide a suitable retaining baffle for the mercury pool. If the movement of the liquid mercury is not sufficiently restricted, it can completely lose contact with the ignitors at times, and can pile up against one side or the other of the tube side wall due to the swaying of the locomotive at high speed. This will not only cause ignitor misfires but many mercury droplets will be thrown up into the top section of the tube, and those which happen to land on the red-hot anode or grids can increase the mercury vapor pressure sufficiently to destroy temporarily the rectifying ability of the tube.

Many configurations of mercury retaining baffles were fabricated and tested. Testing was done by sealing the header-and-baffle assembly into a short piece of cylinder, adding the proper amount of mercury and closing the top of the dummy cylinder securely with a plate of clear plastic so the motion of the liquid mercury could be observed as the test assembly was shaken and vibrated to simulate the motion of a locomotive. Initial tests were made on a shaker table which provided a variety of vertical and horizontal motions, then the most promising designs were mounted in dummy cylinders again and observations of mercury movement were made on locomotives during runs.

Satisfactory baffling is obtained by dividing the cathode pool into small areas

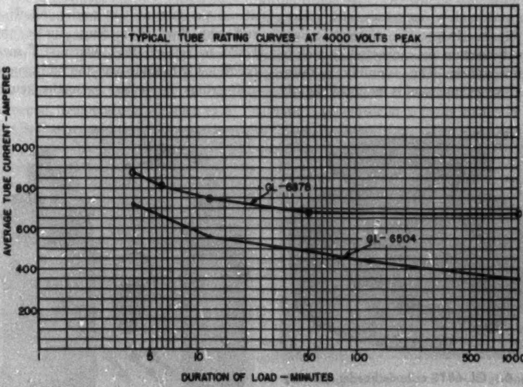


Fig. 4. Typical-tube rating curves at 4,000 volts peak

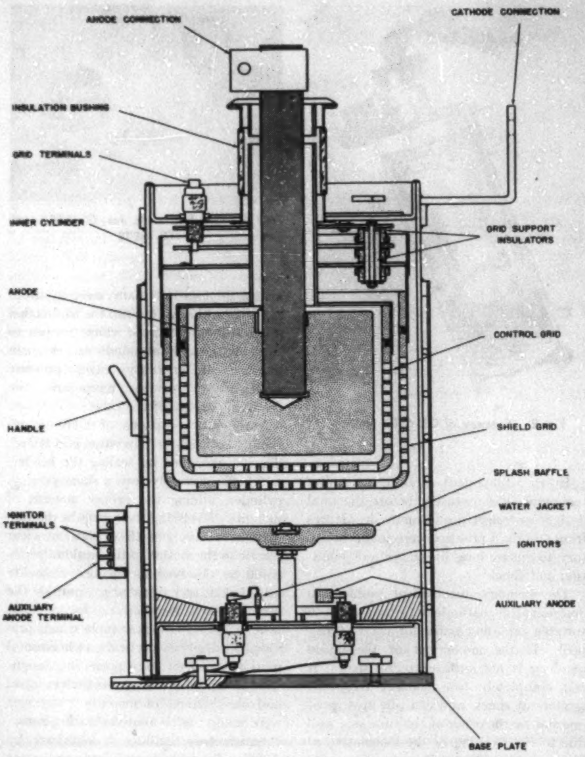


Fig. 5. Cross section of 12-inch locomotive ignitron GL-6878

of approximately one square inch apiece; see Fig. 6. In practice, concentric rings of molybdenum are connected by molybdenum vanes. The assembly is fabricated by welding and is mounted slightly above the face of the header, allowing sufficient flow of mercury under the baffle to replace that which still may occasionally cascade over the barrier and that which is ejected from the center sections by the cathode spot, thereby keeping substantially uniform mercury level at the ignitor.

Water Nipples and Mounting

Space requirements in the locomotive dictated the placement of the water nipples 90 degrees apart and positioned symmetrically opposite the flat surface of the cathode terminal.

The water jacket was made considerably shorter than the inner cylinder on these tubes so that the ends of the heavy-walled inner cylinder could be used as a solid area for clamping the tubes in the rectifier. The tube clamps are supported on shock mounts which prevent transmission of much of the high-frequency component of vibration to the tube structure.

Structural Differences Between GL-6504 and GI-6878 Tubes

There are several structural differences between the 8-inch and 12-inch tubes. The most noticeable being that of diameter, see Fig. 7. Since the 12-inch tubes were designed to handle approximately twice the power of the 8-inch tubes, the diameter of the anode, grids, and cylinders had to be increased. The outside diameters are approximately 13 inches and 9 inches respectively.

As soon as the first 8-inch tubes were installed on locomotives, it was realized that something should be done to facilitate making external electric connections to the ignitor terminals. Accordingly, when the 12-inch *GL-6878* tubes were designed, the ignitor connections were brought out to a terminal block permanently mounted on the front side of the tube for easy access. As part of this same change a steel base plate was bolted to the bottom of the tube to serve as a mounting plate and also to protect the ignitor seals from accidental damage.

The anode connection on the 12-inch tube was changed from a flexible copper cable to two solid copper flags with holes for attaching a separate lead. In this way, the flexible cable can easily be replaced if it should become damaged.

Operating Experience

The 10-passenger locomotives using the *GL-6504* ignitrons were placed in regular service during the period from March 24, 1955 to July 14, 1955. By the time all ten locomotives had been used a year, the average distance traveled by each was well over 100,000 miles. Experience

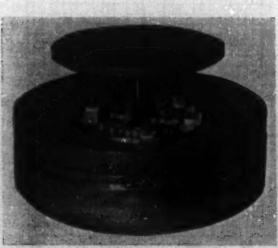


Fig. 6. GL-6878 cathode header assembly

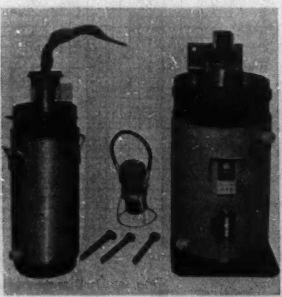


Fig. 7. GL-6504 and GL-6878 ignitrons

indicates that the tube design provides adequate margin for reliable operation in this service. Only 5% of the original tubes have had to be removed because of tube defects during the first $1^1/_2$ years of operation. The 12 3,300-hp freight locomotives have recently gone into service.

Summary

The *GL-6504* and the *GL-6878* ignitrons have been designed electrically and mechanically for application in railway locomotives.

Factory tests and field operating experience prove that ignitrons can be successfully used in rectifier locomotives.

Ram-Air Cooling Systems for Aircraft Generators

RICHARD M. MORONEY
ASSOCIATE MEMBER AIEE

WHEN AIR is taken aboard an aircraft in flight, it undergoes an increase in pressure which is called the "ram effect," and a cooling system which uses such air is called a ram-air cooling system. Up to the present, almost all systems for cooling aircraft generators have been of the ram-air type and, in addition, the simplest subdivision of the type, the so-called blast-cooling system. In this system, the generator is in effect simply hung overboard, and cooled by the stream of air forced through it by ram pressure. (In practice, the generator is mounted inside the aircraft diffuser.)

The effectiveness of blast cooling at low flight speeds is dramatically demonstrated by the difference in size between present-day generating equipment for airborne use and equivalent ground equipment. Unfortunately, there is a temperature rise associated with the ram effect, and this limits the flight speed at which blast-cooling is possible. Proximity to this limit was recognized recently, and curves such as that of Fig. 1 appeared in the literature.[1,2] This curve shows the envelope of flight conditions within which blast cooling is possible, plotted on the Mach number-altitude plane. The left boundary of the envelope does not represent a blast-cooling limitation, but rather a "flyability" limitation. Present aircraft do not fly in the region of the plane to the left of this boundary.

The difference between the curve of Fig. 1 and those of previous reporters is that Fig. 1 is based on the class *C* generator,[3] which was specifically designed to increase the range of blast-cooling operability within the present state of the art of high-temperature generator design. It is seen from Fig. 1 that present fighter aircraft are already at the class *C* blast-cooling limit. Actually, the future of blast cooling is even grimmer than Fig. 1 would lead one to believe. It has been proposed to build a class *D* machine having in effect a temperature capability of 100 C (degrees centigrade) higher than the class *C* machine.[4] Fig. 2 shows the blast-cooling envelope for such a machine, and comparison with Fig. 1 will show how little would be gained by this marked improvement in generator temperature capabilities. The reason for this small gain is that, in the nature of the ram phenomenon, temperatures at inlet to a blast-cooled generator rise in proportion to the square of the flight speed.

In view of these considerations, the need for a move to some new cooling system has been recognized. In this regard, the development of liquid-cooled generators[5] represents a contribution which permits greater flexibility in the choice of cooling method to be employed. It does not appear to be generally understood, however, that liquid cooling does not represent a solution in itself, since a system must be provided to extract heat from the circulating coolant. Thus it is easy for the electric-system designer to specify a liquid-cooled generator with coolant piped to an aircraft's oil, fuel, or hydraulic system, but somebody must worry about cooling these systems, and his problems are already acute.[6,7]

Design of Ram-Air Cooling Systems

Among the new cooling systems that come into consideration are those which continue to use ram air, but which lower its temperature before introducing it into the generator. Thus systems such as that shown in Fig. 3 have appeared in the literature.[8] Systems of this type are feasible because of the high pressure associated with stagnation air that is too hot for direct application as a coolant. This high-pressure high-temperature air may be expanded in a turbine to get low-temperature air which still has sufficient pressure to force itself through generator-cooling passages. Air turbines are highly developed because of their use in cabin-cooling systems.

A turbine requires a load against which to operate, since mechanical power is produced. Lightweight turbines of the type which would be considered for use in aircraft operate at such high rotational speeds that this power could not be put to any use outside of the cooling system, so it seems reasonable to load the turbine with a compressor. As will be seen later, this decreases system drag.

Suppose then that one has made the decision to use compressors as loads for turbines, and it is necessary to design a ram-air cooling system capable of cooling a class *C* generator at a flight condition of Mach 3, 70,000 feet. The following questions arise:

1. Assuming that there is more than one system capable of doing the required cooling job, what shall be the criterion for determination of the optimum system?
2. Given such a criterion, which arrangement of turbines, compressors, and possibly heat exchangers (which "system topology") will contain the optimum system?
3. Within a given topology, which values of turbine expansion ratios, air flows, etc. (which "system parameters") produce the optimum system?
4. Will the optimum system for the given speed and altitude operate satisfactorily at other speeds and altitudes?

Paper 57-472, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted December 27, 1956; made available for printing March 22, 1957.

RICHARD M. MORONEY is with Massachusetts Institute of Technology, Cambridge, Mass.

The author wishes to acknowledge the stimulation and direction provided by A. K. Sen and A. Kusko. Much of the Appendix is an extension of the original work of P. N. Roychoudhury. The research reported was made possible through the support extended the Massachusetts Institute of Technology Servomechanisms Laboratory, by the United States Air Force (Equipment Laboratory, Wright Air Development Center) under Contract No. *AF 33(616) 2921*, M. I. T. Project No. *DSR-7352*. The assistance of Dr. E. O. A. Naumann is gratefully acknowledged.

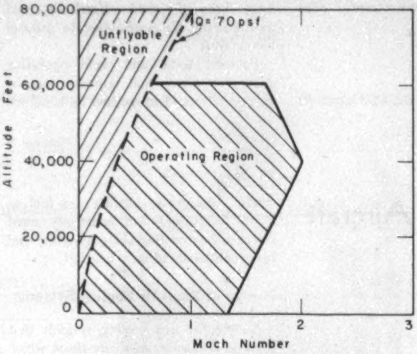


Fig. 1. Flight region of blast-cooled class C generator

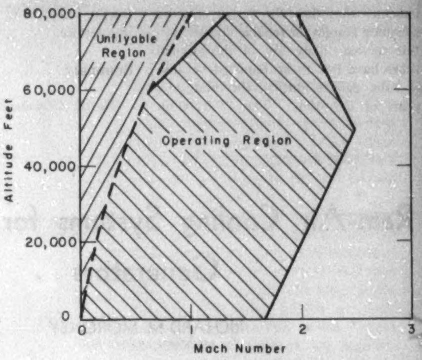


Fig. 2. Flight region of blast-cooled class D generator

The remainder of this section summarizes the results of an investigation of these questions.[9]

OPTIMIZATION CRITERIA

Weight and drag are two factors associated with any ram-air cooling system that detract from the performance of an aircraft in which the system is installed. One pound of drag is in general more harmful than a pound of weight, because under normal flight conditions a pound of engine thrust can be converted into more than a pound of lift, the multiplying factor being the so-called lift-drag ratio L/D. A suitable optimizing criterion thus is

$$\text{Minimum}\left(\frac{L}{D}\times\text{drag}+\text{weight}\right)$$

Note that for certain applications a different criterion might be used. Thus if it is desired to build the fastest-driving aircraft possible, a criterion is

$$\text{Maximum (weight}-\text{drag)}$$

since this will give the largest net thrust in a dive.

SYSTEM TOPOLOGY

Without additional constraints it is difficult to visualize a method of selecting an optimum topology. The approach used here consists of making the additional stipulation that the simplest system possible shall be used, giving the following development; see Fig. 4.

Fig. 4(A) shows the simple blast-cooling system, and (B) shows the turbine inserted to cool incoming air to a usable temperature. Diagram (C) shows a compressor added after the generator, to serve as a load for the turbine and increase the pressure of the cooling air before it is exhausted through the exit nozzle. The system of (C) appears to be the simplest one which may be used to extend the maximum allowable ram air temperature beyond the limit for simple blast cooling.

The flight speed for the system labeled (C) is reached when the turbine expansion ratio necessary to reduce ram temperature to an allowable generator-inlet temperature becomes so large that all of the available ram pressure is used, and there is no pressure drop available across the generator to force an air-flow through it. At this point some additional complication of the cooling system is necessary, and the proper topology change is found by observing that at the point where the system fails to operate, the generator exit temperature is below the ram-air inlet temperature. (The exit air must be below the critical "equivalent wall temperature" of reference 10, and if inlet air were much below this, blast cooling would still be possible.) Evidently generator exit air may be used to pre-cool the air entering the cooling turbine, by use of a heat exchanger as shown in the system labeled (D), giving a "regeneration" effect.

This regenerative system was found to provide adequate cooling at Mach 3 and 70,000 feet, and thus, within the limits of the assumptions made, it represents the optimum cooling system for that design point. It is interesting to observe that the specified flight condition is close to the marginal one for this system, since it was found that a 10% increase in generator loss would require the addition of another turbine, as shown in Fig. 4(E). Fig. 5 is a more descriptive drawing of the system shown in Fig. 4(D).

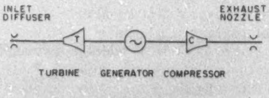


Fig. 3. Simple ram-air cooling system

OPTIMUM PARAMETERS

Optimum parameters within a given topology were found by selecting a number of values for each independent system variable and by programming a digital computer to calculate all the states of air in the system and finally the drag caused by the system. The Appendix presents a complete outline of this method for the regenerative Mach 3 system. Several sets of independent variables which produced low drag systems were examined to determine which would have the lowest weight. Available data on the weight of cooling system components are limited and of unknown accuracy, so the weight calculation was not included in the computer program. The important consideration in minimizing weight appears to be the use of the minimum possible mass flow of air through the system. This is also a consideration in minimizing drag, as is shown by the drag equation in the Appendix. The counterbalancing factors which create an optimizing problem arise when a reduction in system air flow is attempted. To keep the generator cool, inlet temperatures must be reduced (because of the lower heat transfer

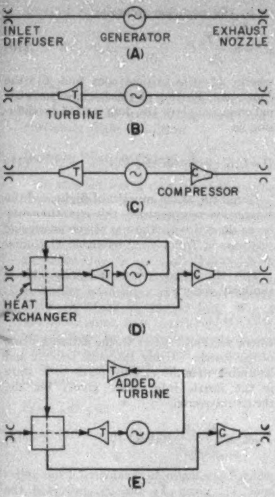


Fig. 4. Development of system topology

coefficient). This calls for larger turbine expansion ratios, which in turn create lower densities and lower system exit pressure. The last two factors increase weight and drag respectively.

The optimum parameters for a 40-kva-generator Mach 3 colling system corresponded to the value of (L/D×drag +weight) of 60 pounds, which is seen to be about equal to the weight of the generator. Thus one may say roughly that electric power costs twice as much at Mach 3 as it does at Mach 1. The saving of weight which might be realized by accepting lower thermodynamic efficiencies in components (these were considered fixed in all computer work) was not evaluated because no data were found relating thermodynamic efficiencies of components to their weights.

OFF-DESIGN OPERATION

The question of off-design operation has not been satisfactorily answered. Some results were obtained based on a set of assumptions about the off-design performance of system components including diffusers and nozzles, and these are presented for academic interest in Table I. The complete results do not give a simple region of operation in the Mach number-altitude plane, whereas such a result seems necessary by intuition. The problem will not be solved until data are available on the performance of turbines, compressors, heat exchangers, and particularly diffusers, when they are operating at other than their design condition.

Table I. Off-Design Operation of Cooling System Designed for Mach 3, 70,000 Feet

Mach No.	System Will Operate, Altitude	System Will Not Operate, Altitude
2.67	60,000	80,000
3.00	20,000	
3.33	60,000	

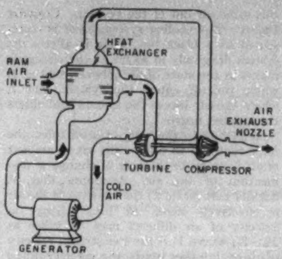


Fig. 5. Functional diagram of regenerative ram-air cooling system

Characteristics of Ram-Air Cooling Systems

Ram-air cooling systems may be used to cool generators at flight speeds up to approximately Mach 3 without requiring a complicated topology. This follows from the preceding section.

Ram-air cooling systems are relatively insensitive to variations in the amount of loss to be dissipated. In support of this statement, Table II shows the change in penalty of a system designed for Mach 3 at 70,000 feet, when the generator of nominal 7.2-kw loss has its losses increased by the amounts shown. Most cooling systems have the property of an increasing penalty with an increase in cooling load as, for example, in the case of an expendable water coolant system. If package cooling systems are planned, i.e., systems cooling a generator and other components simultaneously, this characteristic of ram-air systems makes them appear particularly valuable.

For short flight times, ram-air systems are undesirable. This is because there is no reduction of penalty associated with a shortening of flight time, whereas there is in other systems, notably those using an expendable coolant.

Components used in ram-air cooling systems exist in a high state of development in so far as weight and efficiency are concerned, but at present there are insufficient data available on these components to permit calculation of off-design operation of systems.

Table II. Change in Cooling System Drag With Increasing Generator Losses Mach 3, 70,000 Feet

Per Cent of Standard Loss	Drag, Pounds
100	7.1
110	7.6
115	8.5
120	9.3
125	10.1

Conclusions

Ram-air cooling systems are among the more promising solutions to the thermal barrier problem. The following material is needed before a full-scale development program can be justified:

1. An analytical method for selecting the optimum topology for a given cooling problem.
2. An analytical equation for the weights of components as functions of their thermodynamic variables.
3. Analytical equations describing the off-design and transient performance of cooling system components. Transient information is necessary in order to determine the stability of system transition from one operating condition to another.

Appendix

This Appendix presents a derivation of all the equations necessary to calculate the behavior of a ram-air cooling system operating with a 40-kva class *C* generator. The mechanical operation of programming these equations on a computer is left to the imagination of the reader.

Development of Generator Equations

Page 14 of reference 3 gives the following equations (equations 1 and 2 are here numbered to correspond with those of the reference):

$$\frac{\Delta P}{27.7}\leq P_a:\quad W=\sqrt{\frac{P_a\Delta P}{P_{a0}\Delta P_0}} \tag{1}$$

$$\frac{\Delta P}{27.7}\geq P_a:$$

$$W=\frac{1}{2}\sqrt{\frac{27.7}{P_{a0}\Delta P_0}}\left(P_a+\frac{\Delta P}{27.7}\right) \tag{1A}$$

$$q=w^{0.78}\,\frac{185-t_t}{185-t_{t0}} \tag{2}$$

where the above symbols are defined as

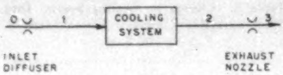


Fig. 6. Ram-air cooling system as seen by engineer calculating drag

follows:

$P_{a0}=14.7$ psia (pounds per square inch absolute)
$\Delta P_0=11$ inches H_2O
$t_{i0}=120$ C
$q=$ratio of losses$=\dfrac{Q}{Q_0}$ $Q_0=415$ Btu/min (per minute)
$w=$ratio of air flows$=\dfrac{W}{W_0}$ $W_0=35$ lb/min
$P_a=$air outlet pressure, psia
$\Delta P=$pressure drop, inches H_2O
$t_i=$air inlet temperature, C

Equation 1 gives the mass flow through the generator, in terms of pressure drop and exhaust pressure, for those cases where the pressure drop across the generator is less than the exhaust pressure, and equation 1A gives the mass flow through the generator for those cases in which the pressure drop across the generator is greater than the exhaust pressure. In all cases, the progress of analysis in this report was such that a generator-inlet pressure and desired mass flow were known, and the generator exhaust pressure was the variable to be calculated. Equation 1 may be rearranged and solved for the exhaust pressure, giving

$$P_a=\frac{P_i+\sqrt{P_i^2-\left(\dfrac{W}{7.24}\right)^2}}{2} \tag{3}$$

where all pressures are converted to psia. and the positive root is selected to make generator inlet and exit pressures equal at zero weight flow. When the quantity under the root sign goes negative, this corresponds to a switch to equation 1A, and the inlet pressure and mass flow are no longer independent.

Equation 2 gives the heat dissipation capability of the generator in terms of the mass flow of cooling air through the machine and the inlet temperature of cooling air. This equation may be solved to find the minimum permissible mass flow for a given inlet temperature

$$W=35\left(\frac{185-120}{185-t_i}\right)^{4/3} \qquad (q=1) \tag{4}$$

For the case of blast cooling, once the stagnation temperature is determined, this equation may be used to find what mass flow the system should be designed to accept, since the lowest possible mass flow will give the lowest drag system. There is no optimization problem in the case of blast cooling.

Calculation of System Drag

The net drag imposed by a ram-air cooling system is the difference between the drag caused by diffusing the air used by the system, and the thrust produced by exhausting this same air out of the system. Consider Fig. 6. The cooling system itself is represented as a black box, since it affects the system drag only in so far as the change it makes in the state of air between that at *1*, where the air enters the system, and *2*, where the air leaves the system and enters the exhaust nozzle.

According to the rules of mechanics, the force required to diffuse the given amount of air is equal to the rate of change of momentum of the air. Assuming that at diffuser exit (state *1*) the velocity of the air is effectively zero, the total change in velocity of air diffused may be taken as V_0-V_1, where V_0 is the free stream velocity of air or, in other words, the velocity of the aircraft. Assuming this velocity constant, the rate of change of momentum is given by this velocity times the rate of change of mass. The latter parameter is in turn given by the relation $\rho A V$ thus giving finally for the drag incurred in the diffusion process

$$D=\rho A V_0^2 \tag{5}$$

to eliminate ρ from this equation substitute M^2C^2 for V^2 where M is Mach number of flight and C is the speed of sound at ambient conditions. The speed of sound may in turn be given by

$$C=\sqrt{\frac{KP}{\rho}} \tag{6}$$

whence

$$C^2=\frac{KP}{\rho} \tag{7}$$

hence

$$V^2=M^2C^2=M^2\frac{KP}{\rho} \tag{8}$$

and

$$\rho A V^2=A M^2KP \tag{9}$$

A is the diffuser cross-section area. To eliminate this A multiply the numerator and the denominator by A^*, using in the numerator Fliegner's equation for A^* as in the equation following

$$A^*=w\frac{\sqrt{T_0}}{P_0}C_1 \tag{10}$$

where $A^*=$the critical area and $C_1=$the constant. Putting all this back in the original equation gives the drag caused by diffusion as

$$D=C_1K\frac{P}{P_0}\frac{A}{A^*}M^2w\sqrt{T_0} \tag{11}$$

Given the Mach number of flight and the stagnation temperature, this equation may be evaluated with the use of the isentropic relations in Table 30 of reference 11 to find A/A^* and P/P_0.

The thrust recovered at nozzle exit is similarly given by

$$T=\rho_3A_3V_3^2 \tag{12}$$

where section 3 refers to the exhaust plane of the nozzle. Using the same scheme and assuming that the exhaust plane total state is the nozzle inlet state gives, for the thrust recovered,

$$T=C_1K\frac{P_3}{P_2}\frac{A_3}{A_3^*}M_3^2w\sqrt{T_2} \tag{13}$$

which may again be evaluated from reference 11, given T_2 and P_2, provided the exhaust plane pressure is assumed to be atmospheric (implying a well-designed nozzle). Combining these results gives for the drag of a ram-air system having a single exhaust nozzle

$$D=KC_1W\left[\left(\frac{P}{P_0}\frac{A}{A^*}M^2\sqrt{T_0}\right)_{\text{diffuser inlet}}-\left(\frac{P}{P_0}\frac{A}{A^*}M^2\sqrt{T_0}\right)_{\substack{\text{nozzle}\\ \text{exit}}}\right] \tag{14}$$

For a given design point, this may be reduced to

$$D=C_3W[C_2-\sqrt{T_i}f(P_a,P_i)] \tag{15}$$

where T_i and P_i are exhaust nozzle inlet conditions and P_a is atmospheric pressure. The function f may be approximated over

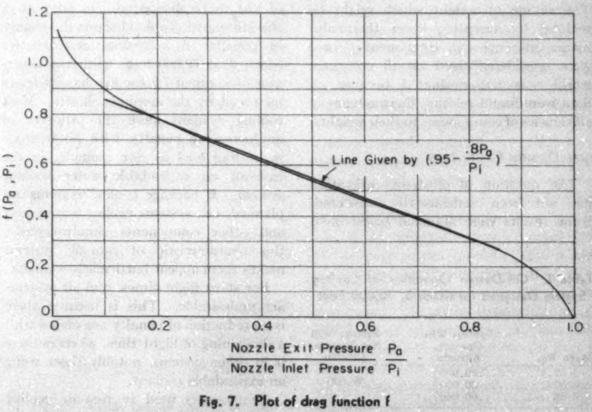


Fig. 7. Plot of drag function f

most of its range by equation 16

$$f(P_a, P_e) = 0.95 - \frac{0.8P_a}{P_e} \tag{16}$$

A plot of this function is given in Fig. 7. The constants C_3 and C_5 in the drag equation depend on the units being used and the particular flight condition.

Sample System Calculation

In this section, a complete calculation of the air states in a representative ram-air cooling system is presented. Consider the system shown in Fig. 8, which is a functional sketch of the system of Fig. 5. To specify an operating condition, the mass flow through the system is chosen as 15 lb/min, and inlet conditions are to correspond to flight at Mach 3, 70,000 feet. For this operating point, inlet conditions, *0* in Fig. 8, are obtained by calculating isentropic stagnation conditions *00* and correcting the pressure so obtained for diffuser efficiency to obtain

$P_0 = 12.43$ psia

$T_0 = 1176$ F abs (degrees Fahrenheit absolute)

For a mass flow of air of 15 lb/min, the maximum generator inlet temperature, condition *2* in Fig. 8, may be calculated with the use of equation 4 (suitably rearranged and corrected for use of F abs instead of C) to get

$T_2 = 604$ F abs

The generator exit temperature is now found by using the first law in the form

$$T_3 = T_2 + \frac{Q}{wc_p} \tag{17}$$

where Q, the generator losses, is taken as 415 Btu/min (7.3 kw). This gives

$$T_3 = 604 + \frac{415}{15 \times 0.24} = 719 \text{ F abs}$$

The maximum heat exchanger effectiveness which may be used without going to an excessively heavy exchanger is 0.7. With this effectiveness, it is possible to solve for T_1 as follows: By definition, the effectiveness e is

$$e = \frac{T_0 - T_1}{T_0 - T_3} \tag{18}$$

whence

$$0.7 = \frac{1176 - T_1}{1176 - 719}$$

$T_1 = 856$ F abs

Since the temperatures at both sides of the turbine are now known, it is possible to calculate the required turbine pressure ratio, given an assumed turbine efficiency of 0.8. This is done by the use of the perfect gas laws which yield the equation

$$T_2 = T_1\left[1 - \left(1 - \frac{1}{\gamma_1^{0.283}}\right)\eta_T\right] \tag{19}$$

where γ_1 is the turbine pressure ratio expressed as a number greater than 1, i.e., P_1/P_2, and η_T is the turbine efficiency, assumed 0.8. Solving this equation for γ_1 gives

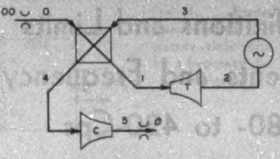


Fig. 8. Schematic diagram of regenerative ram-air cooling system

$$\gamma_1 = \left(\frac{\eta_T}{\frac{T_2}{T_1} - 1 + \eta_T}\right)^{\frac{1}{0.283}} \tag{20}$$

$\gamma_1 = 5.06$

If a 10% pressure drop is assumed in the heat exchanger, P_1 is given by

$$P_1 = 0.9P_0 = 0.9 \times 12.43 = 11.18 \text{ psia} \tag{21}$$

and P_2 may now be calculated since γ_1 is known, giving

$$P_2 = \frac{P_1}{\gamma_1} = \frac{11.18}{5.06} = 2.20 \text{ psia} \tag{22}$$

The pressure drop through the generator is calculated using equation 3 which gives the following

$P_3 = 1.49$ psia

and now, after a second pressure drop in the heat exchanger, also assumed 10%, the pressure at *4* in Fig. 8 is

$$P_4 = 0.9P_3 = 0.9 \times 1.49 = 1.34 \text{ psia} \tag{23}$$

From the first law again, the heat rejected by the air in passing through the heat exchanger from *0* to *1*, Fig. 8, must be absorbed by the air passing in the other direction from *3* to *4*, hence

$$T_0 - T_1 = T_4 - T_3 \tag{24}$$

$$T_4 = T_3 + T_0 - T_1$$

whence

$T_4 = 719 + 1176 - 856 = 1039$ F abs

It is assumed that no mechanical power is added to or extracted from the system, whence the compressor power input must equal the turbine power output. Continuing to assume constant specific heat, this condition may be expressed algebraically as

$$T_5 - T_4 = T_1 - T_2 \tag{25}$$

whence

$T_5 = 1{,}039 + 856 - 604 = 1291$ F abs

If a specific piece of hardware were being used as the cooling system compressor, it might be best considered as a constant-pressure-ratio fan. In the general case, the perfect gas laws may be used as in the turbine calculation to get

$$\gamma_2 = \left[\left(\frac{T_5}{T_4} - 1\right)\eta_c + 1\right]^{\frac{1}{0.283}} \tag{26}$$

where

$$\gamma_2 = \frac{P_5}{P_4}$$

η_c = compressor efficiency—assumed 0.72.

which gives

$\gamma_2 = 1.76$

and hence P_5 is given by

$P_5 = P_4\gamma_2 = 1.34 \times 1.766 = 2.33$ psia

P_5 and T_5 may now be inserted in the drag equation, giving

$$D = 0.0451W\left[38.6 - \sqrt{T_5}\left(0.95 - \frac{0.512}{P_5}\right)\right]$$

$$D = 0.0451 \times 15\left[38.6 - \sqrt{1291}\left(0.95 - \frac{0.512}{2.36}\right)\right]$$

$= 7.1$ lb

References

1. **Equipment Cooling Systems for Aircraft, R. H. Zimmerman, W. Robinson, K. G. Hornung, W. E. Kraus.** *Technical Report No. 54-359, Part C*, Wright Air Development Center, Dayton, Ohio, 1954.

2. **Cooling Problems of Electrical Equipment in High-Speed High-Altitude Aircraft, E. O. A. Naumann.** *Technical Digest*, Armed Services Technical Information Agency, Dayton, Ohio, 1955, Fig. 10.

3. **Establishment of Standard Rating and Cooling Air Flow Requirements for AC Aircraft Generators, E. O. A. Naumann.** *Technical Note WCLE 45-23*, Wright Air Development Center, 1954, 22 pp.

4. **Guide of New Temperature Requirements for Aircraft Electrical Equipment.** *WCLE Disposition Form, Ibid.*, 1957 Program, 4 pp.

5. **An Oil-Cooled A-C Generator for High-Speed High-Altitude Aircraft, H. J. Braun, W. J. Shilling.** *AIEE Transactions*, vol. 74, pt. II, 1955 (Jan. 1956 section), pp. 456–61.

6. **Effect of Supersonic Flight on Power-Plant Installation Systems, R. B. Keusch.** *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 77, July 1955, pp. 721–27.

7. **A Study of Turbine Engine Oil Cooling Problems in High Speed, High Altitude Aircraft, George T. Papadopoulos.** *WADC Technical Report No. 55-172; ASTIA No. 73718* (Confidential), Armed Services Technical Information Agency, Wright Air Development Center, Dayton, Ohio, 1955.

8. **Temperature Problems of Equipment in High-Speed Aircraft, H. W. Adams.** *Transactions*, American Society of Mechanical Engineers, vol. 77, July 1955, pp. 735–41.

9. **Cooling and Materials Investigation for Aircraft Generators, A. Kusko, et al.** *Technical Report No. 56-305*, Wright Air Development Center, 1956.

10. **Correlation of Rating Data for Rotating Electric Machinery, Erwin O. A. Naumann.** *AIEE Transactions*, vol. 73, pt. II, 1954 (Jan. 1955 section), pp. 443–52.

11. **Gas Tables (book), J. H. Keenan, J. Kaye.** John Wiley & Sons, Inc., New York, N. Y., 1948.

Development of Definitions and Limits for Frequency Transients and Frequency Modulation in 380- to 420-Cps Aircraft Electric Systems

OSCAR MARKOWITZ
MEMBER AIEE

Synopsis: This paper describes in detail the work of the Naval Air Development Center (NADEVCEN) which attempted to develop the definitions and limits for frequency modulation and frequency transients in 380- to 420-cps (cycles per second) aircraft electric systems. In order to provide a common base for such definitions, the instruments used for making the measurements of these characteristics were examined and delineated in terms of required performance criteria. The definitions and limits developed are recommended for integration into appropriate specifications.

IN THE DESIGN of present utilization equipments for aircraft with constant-frequency electric systems, transient-frequency characteristics of the electric system are not taken into consideration, because these characteristics have not been defined nor have their limits been specified in the appropriate military specifications. Thus, the equipment often fails to give anticipated performance in the aircraft, even though it provided full performance under laboratory conditions. This has forced industry to set many arbitrary definitions and/or limits for transient frequency, but to date there has been no uniformity and little agreement as to the extent, rate, or duration of the transient-frequency changes acceptable for aircraft constant-frequency electric systems.

Confusion in the setting of limits for the transient-frequency characteristic has been created by the lack of commercially available instrumentation to measure frequency on a transient basis. Each engineering group interested in measuring transient frequency has used its own instruments with little intercorrespondence.[1] This variation in instruments has been unavoidable, since the problem has not been sufficiently defined to determine exactly what information is desired from the instrumentation.

Few data are available on the characteristics of frequency modulation within constant-frequency aircraft electric systems. Recent compatibility problems in the field have warned the designers of electric systems that frequency modulaton must be contained within definite limits, and the designers of utilization equipment must consider the presence of frequency modulation in their designs. However, compatibility cannot be assured until characteristics are defined and limits have been established.

Among the weaknesses of Specification *MIL-E-7894*[2] revealed by a NADEVCEN study, was the absence of definitions and limits for frequency transients and frequency modulation. This paper discusses the problem and requirements for adequate instrumentation, and establishes definitions and limits for the transient-frequency and frequency-modulation characteristics of constant-frequency aircraft electric systems.

Discussion of the Problem

In aircraft electric-power systems of the so-called constant-frequency types, the frequency does not remain constant, but varies during quiescent operation within some frequency band, such as 380 to 420 cps, called the steady-state limit. Sudden changes in frequency within and beyond this band can occur during electric-system load disturbances or drive-system input disturbances. When the disturbances are transient in nature, corresponding frequency variations, including the return to steady-state limits due to regulation controls, are called transient-frequency characteristics. Until the characteristics of all frequency variations are defined within specified limits, designers of utilization equipment cannot assure compatibility of their equipment with aircraft electric systems.

In accordance with general usage, frequency variations fall within the following three broad categories, classified by virtue of the extent, rate, and duration of frequency changes.

1. Transient-frequency changes, which imply a sudden change in frequency within the definition of transient, deviate greatly and with high rates of frequency change. The deviation and return to steady state occurs within a short time, perhaps one second.
2. Frequency modulation is the cyclic and random, continuous or dynamic, variation of frequency about a mean frequency during steady-state electric system operation. The frequency modulation is normally within narrow frequency limits, and occurs as a result of speed variations in the generator rotor which in turn are due to dynamic operation of drive and rotor coupling. Rates of frequency changes in frequency modulation are much lower than those caused by transient frequency changes. The mean frequency can change as a result of frequency drift.
3. Frequency drift is the extremely slow and random variation in frequency within the steady-state limits which are caused by environmental or service effects on the electric-power drive system. The rates of frequency change are much lower than those of other frequency variations.

This paper deals with the transient frequency and frequency modulation defined in items 1 and 2 of the foregoing.

Many utilization equipment designers have indicated that their equipment is affected by the degree of transient-frequency excursion, and also by the rate of frequency change during the excursion. Thus, it becomes important to examine this rate characteristic of frequency transients, and contain it within specified limits. Since information about rate of frequency change has not been required from the instruments providing transient-frequency recordings, such instruments vary widely in their ability to show high rates of frequency change.

Many data have been published showing oscillograms of frequency transients in constant-frequency systems. The data, when assembled and integrated, could provide a basis for the preparation of defined limits for the transient-frequency characteristics. However, the data can-

---

Paper 57-482, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7-9, 1957. Manuscript submitted August 20, 1956; made available for printing February 21, 1957.

OSCAR MARKOWITZ is with the Naval Air Development Center, Johnsville, Pa.

Acknowledgment is made to the many engineering organizations involved in aircraft electric systems who generously provided NADEVCEN data for its study of frequency transients and frequency modulation. The mathematical analysis in Appendixes I and II were made by Samuel Wolin of NADEVCEN. The Naval Air Material Center originally designed the instrumentation in Fig. 3. It was redesigned by NADEVCEN.

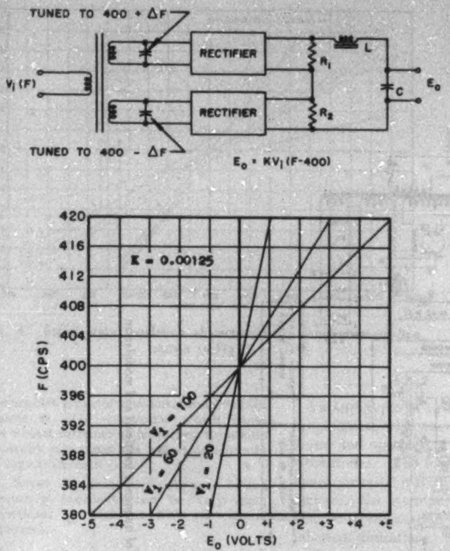


Fig. 1. Fundamental voltage frequency relationships for a frequency discriminator

not be considered valid until the instrumentation for each set of data is verified in the light of projected definitions and limits for the frequency transients.

There are not enough accumulated or published data on frequency modulation to establish limits of frequency modulation or definite design criteria which could be used with component and system design. Limited tests to obtain data of frequency modulation were made at NADEVCEN to establish a starting base for developing definitions and limits.

Instrumentation

Instrumentation for frequency transients is generally accomplished by using a voltage analog of frequency deviation from a given frequency. The voltage analog is then fed to a recorder having satisfactory response time. Since the frequency transients in electric systems are caused by disturbances which simultaneously create electric-system voltage transients, the frequency analog must be insensitive to the electric-system voltage changes over extremely wide limits, even though the electric-system voltage is used to provide the frequency information. None of the various instruments used in the past is sufficiently insensitive to electric-system voltage change to give any accurate indication of the frequency changes for all conditions of electric-system operation.

Another method of measuring frequency deviations is making direct tachometer connection to the generator drive equipment. The tachometer voltage output is a function of drive speed, allowing the output voltage to be calibrated in terms of frequency. This method cannot be used universally, since tachometers cannot be mounted easily on drives or generators except with some laboratory operations. This method also has other disadvantages, in that data involving small frequency deviations are difficult to obtain, and generator transient-flux shifts and shaft torsionals are not seen. Because of these inherent limitations, this type of instrumentation was not considered in this project investigation.

Because of the diversity of instruments used to measure frequency transients and frequency modulation, the obtainable data have not been amenable to correlation. It is hoped that when the variations and inaccuracies of presently used instrumentation are brought to light, and the actual requirements for all future instruments as a common denominator have been specified there will be consistent data with sufficient accuracy for correlation.

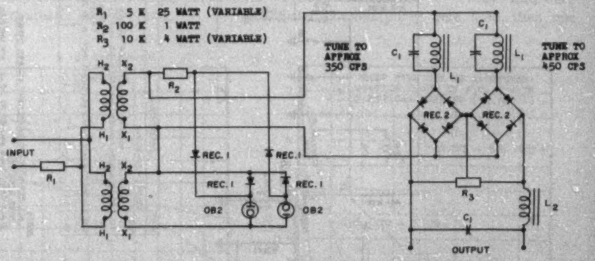
The method of measuring frequency transients has been to use a frequency discriminator centered at 400 cps to obtain a voltage analog of the frequency change from 400 cps. There has been sufficient experience in the use of frequency discriminators, in general, and for 400 cps in particular, to show that it is not difficult to obtain discriminators with good linearity through reasonable frequency changes. However, the discriminator by itself cannot provide the complete picture of transient frequency changes for aircraft a-c electric systems.

In examining composite characteristics of aircraft constant-speed electric systems, it is found that:

1. The steady-state operating frequency is usually near 400 cps, and may be anywhere in the band between 380 to 420 cps.

2. Frequency transients may travel for the most extreme conditions to 320 or to 480 cps. This represents ±20% change from a 400-cps base.

3. Disturbances which cause transient



Courtesy of Westinghouse Electric Corporation

Fig. 2. Frequency-sensing circuit

Transformer—1 to 4
C_1—0.25 reference
L_1—UTC VI-C 11
L_2—UTC VI-C 19
Recorder 1—Fansteel, BD551T
Recorder 2—Fansteel, BB066T

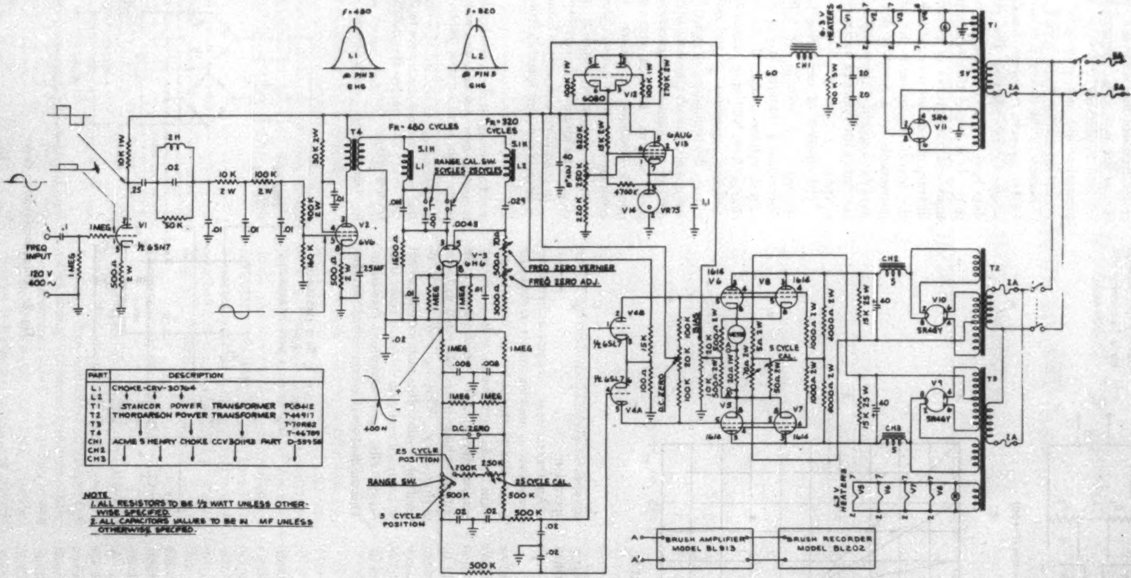

f = 480
f = 820
L1
L2
@ PIN 3
6H6
FREQ INPUT
120 V
400 ~
1MEG
½ 6SN7
V1
10 K 1W
.25
2 H
.02
50 K
10 K 2 W
100 K 2 W
.01
400 K 2 W
140 K
V2
6V6
25 MF
500 Ω
30 K 2 W
T4
FR = 480 CYCLES
FR = 320 CYCLES
5.1 H
RANGE CAL. SW.
5 CYCLES 25 CYCLES
.0048
.029
.011
.001
V-3
6H6
1500 Ω
1MEG
.01
20 Ω
500 Ω
3000 Ω
FREQ ZERO VERNIER
FREQ ZERO ADJ.
.02
.008
D.C. ZERO
400 N
25 CYCLE POSITION
5 CYCLE POSITION
RANGE SW.
100 K
250 K
25 CYCLE CAL.
500 K
V12
6080
100 K 1W
270 K 2W
15 K 2W
820 K
250 K
220 K
6AU6
V13
4700 K
V14
VR75
CH1
100 K 5W
20
40
5V
SR4
V11
T1
6.3 V HEATERS
2A
8A
V4B
V4A
½ 6SL7
15 K
100 Ω
D.C. ZERO
100 K
20 K
BIAS
1614
V5
V6
V7
V8
5 CYCLE CAL.
4000 Ω 2 W
1000 Ω 2 W
CH2
CH3
15 K 25 W
V9
V10
5R4GY
T2
T3
6.3 V HEATERS
A
A'
BRUSH AMPLIFIER MODEL BL 913
BRUSH RECORDER MODEL BL202
PART DESCRIPTION
L1 CHOKE-CRV-30764
L2
T1 STANCOR POWER TRANSFORMER PC8412
T2 THORDARSON POWER TRANSFORMER T-44917
T3 T-70R82
T4 T-46789
CH1 ACME 5 HENRY CHOKE CCV30112 PART D-59956
CH2
CH3
NOTE
1. ALL RESISTORS TO BE ½ WATT UNLESS OTHERWISE SPECIFIED.
2. ALL CAPACITORS VALUES TO BE IN MF UNLESS OTHERWISE SPECIFIED.

Fig. 3. Diagram of instrumentation for the measurement of frequency transients and modulation

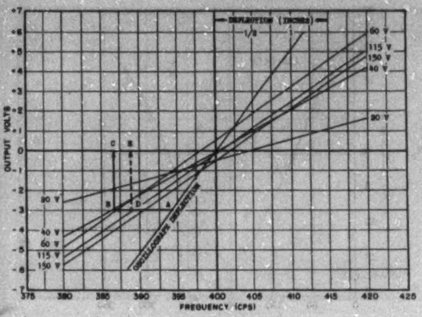


Fig. 4. Steady-state regulation characteristic for the instrumentation shown in Fig. 2

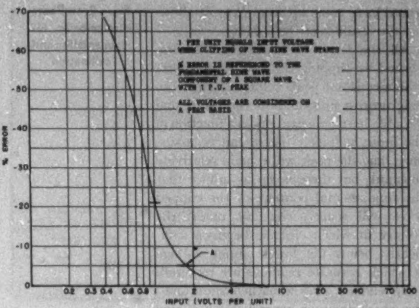


Fig. 6. Fundamental error introduced in clipping a sine wave

frequencies also cause simultaneous voltage transients. The conditions which cause the widest extremes in frequency transients generally also cause the widest extremes in voltage transients.

4. Rates of frequency change during frequency transients can be very high, sometimes as high as 500 per second (cps/sec).

These are the system characteristics which should be used as the basis for instrumentation requirements in the measurement of the frequency transients.

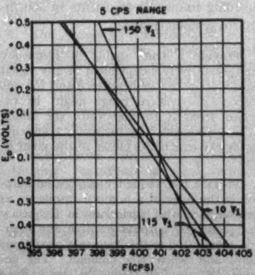
During extreme faults in the electric system, system voltage available for the measurement of frequency can drop to as low as 5 volts, and during recovery of the fault the voltage transient can rise as high as 250 volts. These are line-to-neutral rms values. Any frequency-measuring instrument, which uses line voltage as input, should maintain its output proportional to frequency change, while remaining unaffected by the wide input voltage swings. This is a severe requirement not completely met by known instrumentation.

Various types of instrumentation were investigated, with the main types reproduced and operated in the NADEVCEN laboratories. The investigation revealed many areas of inaccuracies and limiting factors. No attempt was made to develop instrumentation to overcome all the inherent limitations.

The usual varieties of basic-frequency discriminators have an output which is proportional to both input voltage and input frequency. Fig. 1 shows the basic relationships. It is apparent, when the input voltage varies while the frequency varies, that interpretation of the output becomes difficult and inaccurate. In addition, when the frequency varies at rates which cannot be followed by the instrumentation, interpretation of the true frequency variation from the output becomes impracticable. In order to maintain the output of the frequency discrimintor as a true analog of frequency variation, there must be some form of voltage regulation or limiting between the discriminator and its input.

Fig. 2 shows the circuit of a unit which uses voltage regulation characteristics of gas discharge tubes to maintain a constant voltage to the discriminator. Fig. 3 is a diagram of a unit which uses electronic limiting to maintain constant voltage to the discriminator. The range of voltage regulation of the first unit (Fig. 2) on a steady-state basis is from 54 volts to 150 volts. The voltage-frequency characteristics are shown in Fig. 4. The voltage limiting of the second unit (Fig. 3) is fairly good down to 10 volts on the ±5-cycle span, but is poor for the ±25-cycle span. Its characteristics are illustrated in Fig. 5.

The use of Fig. 4 can be illustrated by assuming a fault test, where frequency is measured when the line voltage has been reduced to 60 volts. Assuming an oscillograph deflection of negative 0.6 inch, corresponding to point *A* on the deflection curve of Fig. 4, one moves directly across left as far as the 60-volt curve to point *B*, and then up to 386.75 cps, point *C*. If the instrumentation were assumed to be independent of voltage, and only the 115-volt curve were used, then one



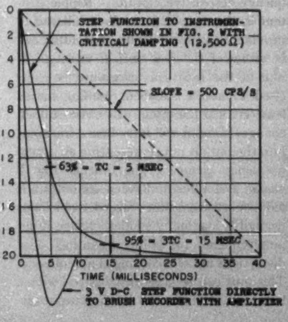


Fig. 5 (above). Steady-state regulation characteristic for the instrumentation shown in Fig. 3

Fig. 7 (right). Transient characteristics of the instrumentation shown in Fig. 2

would have come up from D to E, allowing the result to err by 2.25 cps.

All the forms of limiting are accomplished by clipping the sine wave to provide a constant peak voltage. This clipping process introduces a very large and variable harmonic content as a function of the input voltage. There can be assurance that the peak will remain relatively constant, but there is no assurance that the fundamental content of the clipped wave will remain constant unless the voltage at which clipping begins is sufficiently below the lowest input voltage at which accurate readings are required.

Appendix I gives a mathematical analysis which derives a generalization of sine wave component for the first term of a Fourier analysis. This generalization shows the varying first term as a function of the varying input voltage E_M with the clipping level remaining constant. Fig. 6 plots error from the derived generalization:

$$A_1=\frac{2E_M}{\pi}\text{ arc sin}\left(\frac{b}{E_M}\right)+\frac{2b\sqrt{E_M{}^2-b^2}}{\pi E_M}$$

above 1 per unit (pu $=b/E_M$). Below 1 pu there is no clipping and A_1 follows E_M down linearly to the lowest point considered (0.4 pu), also shown in Fig. 6.

For example, if 5% were allowed as maximum error, and 20 volts peak considered the lowest voltage, Fig. 6 could be applied as follows:

1. Point A is the 5% error level, and correlates to 1.8 pu.
2. Since 1.8 pu represents the lowest voltage level for 5%, it corresponds to the 20 volts. Therefore, 1 pu equals 20 divided by 1.8, or 11.1 volts.
3. The resultant 11.1 volts is the voltage at which clipping would have to begin to maintain a 95% accuracy (factor due to clipping only) at the assumed low-voltage operating point.

The instrumentation discriminator outputs are rectified a-c from the frequency responsive circuits. Ripple content in this output generally is filtered to provide clean d-c output. The filter can be either of the R-C or L-C variety. If it is of the L-C type, care must be exercised to assure sufficient loading for at least critical damping, otherwise sudden changes in input will introduce a transient into the output not representing an actual frequency change. The unit in Fig. 2 has an L-C filter and when its output is fed to a brush amplifier-recorder which has negligible loading, a sudden change in input frequency (or voltage) can introduce distortions in the recording of a frequency. This unit can be critically damped with a load of 12,500 ohms. In the output circuits which have R-C filters, critical damping is not a consideration.

Dynamic characteristics of the frequency discriminator are controlled to a large degree by the type and time constant of its output filter. Since the output is used with some recording means to obtain a record of frequency transients, then the frequency discriminator dynamic action considerations should also include the time constant of the recorder. Consideration of the dynamic characteristic is important, since it will determine the maximum rate of frequency change that the instrumentation can follow accurately. The voltage analog of frequency, after being affected by the afore-mentioned time constants, should be capable of indicating maximum rates of frequency change of at least 500 cps/sec with an accuracy better than 95%. This can be interpreted to mean that the slope obtained when using three times the time constant of the instrumentation as the ordinate and given voltage output in terms of its frequency analog, as the abscissa will define the maximum rate of frequency change that can be reproduced within an accuracy of 95%. Fig. 7 illustrates such considerations with a frequency scale spread that was used with the discriminator of Fig. 2 when in combination with a brush amplifier model *BL-913*, and brush recorder model *BL-202*. Reproduced rates of frequency change of 500 cps/sec were within the 95% accuracy. It is apparent that by increasing the frequency scale span, higher rates of frequency change can be reproduced with a given set of equipment. This is consistent with the general requirements, since the larger rates of frequency changes are experienced during the transients, which also result in the largest frequency swings. The instrumentation shown in Fig. 3 has a discriminator output filter which does not allow for accurate reproduction at the required high rates.

To summarize, instrumentation for recording frequency transients in 380- to 420-cps systems should:

1. Provide a true analog of frequency change relatively independent of any signal input voltage change between 5 and 250 rms volts. This voltage range might be extended from 3 to 250 volts for added facility, while checking frequency transients during extreme faults closer to the point of fault.
2. Provide the frequency change information in the form of a nominal frequency as zero output, with variations in frequency represented by outputs positive and negative, correlating to the frequency variations above and below the nominal. The accuracy of any final frequency steady-state deviation from nominal should be within 2% of full-scale reading. For facility in

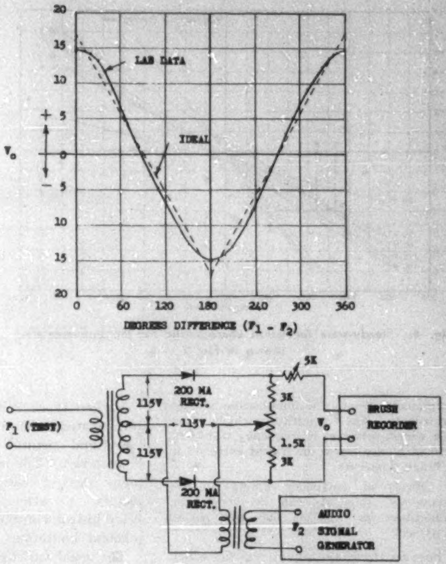


Fig. 8. Phase discriminator instrumentation and calibration

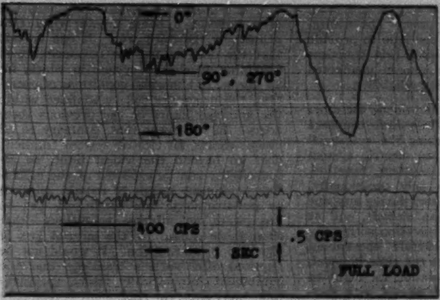


Fig. 9 (above left). Recording of frequency modulation with Navy Type E-1614-1, 2,500-volt-ampere inverter

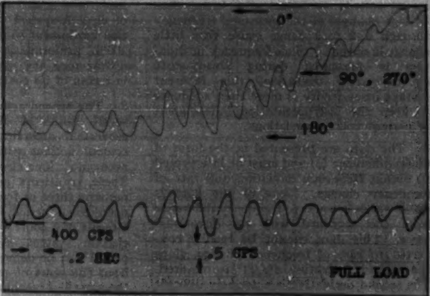


Fig. 11 (above). Recording of frequency modulation with a hydraulic constant speed drive to a 40-kva generator

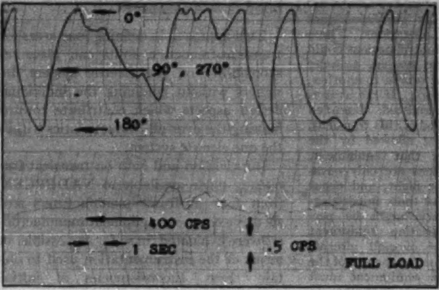


Fig. 10 (left). Recording of frequency modulation with Navy Type E-1737, 1,500-volt-ampere inverter

reading a recording, there can be more than one range in the calibrated frequency spans. Suggested ranges are ±4 cps, ±20 cps, and ±80 cps. It would also be helpful in the use of such instrumentation if the center frequency, which is nominally considered as 400 cps, could be readily shifted in a calibrated manner within the band between 380 and 420 cps, while maintaining any given frequency span.

3. Be capable of reproducing rates of frequency change of 500 cps/sec with an accuracy of at least 95%. It would be acceptable, for any full-scale span of ±10 cps or less, to have an accuracy within 95% with rates of frequency change of 250 cps/sec.

4. The frequency discriminator portion of the instrumentation should be capable of a ±5-volt output for each frequency span with an internal impedance of not more than 5,000 ohms. This would allow reasonable facility in matching the output to most commercial oscillographs. There should be assurance that this part of the instrumentation will have at least critical damping regardless of what recording means are used.

If all instrumentation for measuring frequency transients can meet these performance characteristics, in the future the lack of correlation between data will be rectified.

Measurement of frequency modulation can be accomplished by the use of the equipment previously discussed. This equipment has its limitations for this purpose, in that the frequency modulation for all normal conditions will be only small frequency deviations, and will represent only 5 to 25% of full-scale deflection on a sensitive range as ±4 cycles. Thus, the same reasonable levels of interference which can be present during measurements of frequency transients could blank useful information while making measurements for frequency modulation. If instrumentation design can allow increasing span sensitivity, without correspondingly increasing the minimum levels of interference, and also allow avoidance of "ringing" with increase of Q in the discriminator-tuned networks, then a span of ±2 cycles would be most adequate for direct measurement of frequency modulation.

Another approach in instrumentation for measurement of frequency modulation is to obtain a voltage analog of phase difference between the measured frequency and an arbitrary frequency standard, which should be maintained accurately and conveniently to within 1/20 cps of the mean of the measured frequency for the duration of the measurement. This allows full-scale operation of the instrumentation for extremely small values of frequency modulation, and interference levels are very low because of inherent rejection of noise by a phase demodulator. This scheme is compromised in flexibility of measurement when the standard is a fixed frequency which cannot be varied, or a variable frequency which cannot be varied in small enough increments. Phase discriminators or demodulators of various types[3,4] can be applied to this approach in instrumentation. One arrangement, as set up in the NADEVCEN laboratories, and its calibration are shown in Fig. 8. A phase shifter used for the calibration is described in reference 5. Recordings made with this instrumentation are shown in Figs. 9 through 11.

Other considerations with this type of instrumentation are as follows:

1. Direction sense of measured frequency is lost, since output is plus or minus only in accordance with the relative instantaneous phase difference between measured and standard frequency. There is no direct, calibrated way of knowing if the measured frequency is above or below the standard frequency.

2. There is an inherent degree of voltage insensitivity. If the standard frequency

input voltage is sufficiently high, then voltage changes of the measured frequency through a 2 to 1 range cause very little change in output. Since frequency modulation is measured during steady-state electric-system operation, this inherent voltage insensitivity is more than adequate to keep the instrumentation independent of normal voltage variations.

3. The data are presented in the form of phase difference (ϕ) and must be interpreted to obtain frequency deviation and rate of frequency change. Slope of the data is $d\phi/dt$ which is equal to frequency difference from the standard with no directional sense. This data cannot be readily converted to rate of frequency change, since the second derivative $d^2\phi/dt^2$ is required. The second derivative is extremely difficult to obtain experimentally from the complex wave of phase deviation normally obtained. Some engineering groups have taken frequency-modulation data simultaneously with both the frequency-discriminator and phase-discriminator types of instrumentation; the phase discriminator to give accurately small frequency differences, and the frequency discriminator to give sense of frequency change, rate of frequency change, and the larger frequency deviations.

Frequency Transients

Transients within aircraft electric systems are considered, by utilization equipment designers, on the basis of the extent and rate of their excursions and reoccurrence within one equipment or aircraft operating period. Different design considerations can be made for those transients which occur rarely, and for those which occur often. Possible design correlation between the aircraft electric system and utilization equipment becomes even more apparent with the consideration that the extreme transients are those which occur rarely, and only as a result of faulty electric-system operation, while the smaller transients can occur often as a result of all the normal operations of the electric system. Between the two conditions of rare and normal operations, can be classed as a group the occasional systems operations which result in large but not extreme transients.

Definitions of the electric-system transient characteristics could be made on this correlated basis, between actual electric-systems operation and utilization design considerations. Thus, the transient limits can be defined as those contained within these areas:

1. The first area would define the electric-system transient limits which are a result of all normal and usual aircraft electric-system operations, during flight preparation, air-borne conditions, and landing, which are required for function of aircraft and aircraft mission. Thus, a transient within this area is obtained as a result of normal connection and disconnection of loads, and could occur at any given instant of electric-system operation. The utilization equipment should continue to provide 100% performance during this transient disturbance and should remain unaffected in terms of its continuing reliability.

2. The second area would define the more severe limits of electric-system transient excursions occurring during the occasional control actions within an electric system necessary for electric system function. These transients occur only occasionally, perhaps three or four times during any one operational period for an aircraft, and as a result of a controlled operation within the electric system. An aircraft operational period can be defined as the time interval, from the start of preparation for flight to the postflight engine shutdown with consequent deactivation of the electric system. These larger electric-system transients are usually characterized by correspondence with specific aircraft operations. These normal electric-system control actions could be the start of electric system, transfer of a power bus from one source to another, or paralleling of generators. If the utilization equipment is permitted degraded performance for the period of this transient, it is only to the extent allowed by the detail specification for that equipment. Since these control actions are only occasional during any single flight, and occur for only short periods of time, minor degradations of performance might be permitted of certain utilization equipment without compromise of over-all aircraft performance. In any case, by the end of the transient, the utilization equipment must have fully recovered automatically to 100% performance with unimpaired reliability.

3. The third area would define the extremes of electric-system transients which result from unexpected rather than controlled electric-system actions. When the transient excursion does go into this area, it is the result of fault conditions within the electric system. These transients occur rarely, perhaps once during a flight, or perhaps never during the life of an aircraft; however, when they do occur the exact moment of occurrence is not usually anticipated. These faults could be of a sufficiently serious nature to compromise the electric system during the period of the fault, and therefore utilization equipment could also be permitted correspondingly momentary loss of performance. For this area of transients, correlation with utilization equipment could mean that:

(*a*). If any performance is required, it should be specified in the equipment detail specifications.

(*b*). After the transient, automatic return to 100% performance may not be required by the detail specification.

(*c*). Effect on reliability should be negligible.

(*d*). No unsafe condition should arise.

When the fault is cleared, the electric system should go back to normal, steady-state operation through automatic or manual controls.

4. For the sake of exactness and full delegation of responsibility, there should also be anticipated the unlikely condition of a transient's being greater than that defined by the third area. It can be said that the designer has the responsibility of providing an electric system in which, for all normal or abnormal conditions of operation, the transients will be confined within the specified limits of the previous three areas. In the event that, contrary to all design efforts, an electric system transient does go beyond the widest limit, no responsibility can be expected of the utilization equipment except to remain as much as possible in a safe condition; i.e., to precipitate no adverse conditions other than those caused by the natural loss of the equipment function. It would not be required after such a transient that the equipment automatically return to operation; nor, when the equipment is manually returned to normal operation, that reliability remain unimpaired.

This approach in defining four areas for transient-frequency limits can place responsibility on the electric system designers without confining the individual design aspects which contribute toward making the over-all characteristics within the one electric system.

Information and data on transient frequency made available to NADEVCEN from four air-frame manufacturers and six drive or power equipment manufacturers were evaluated, in so far as possible, in terms of the instrumentation used to obtain them. Inconsistencies of initial rates of frequency changes during frequency transients indicated the need to examine closely the instrumentation used to obtain the data. In general, for frequency transients involving simultaneous heavy-voltage swings, the data could not be considered accurate.

Data on a limited number of types of generators, regulators, drives, and inverters were obtained in the NADEVCEN laboratories for both frequency transients and frequency modulation, for the purpose of supplementing the data received from the other engineering groups.

The curves in Fig. 12 were developed to define the over-all characteristic of frequency transients for 380- to 420-cps electric systems in terms of the limits of frequency excursions as a function of time.

It was attempted to have this over-all characteristic consider the following types of constant speed drives:

1. Hydraulic.
2. Air bleed.
3. Air bleed and burn.
4. Gas turbine.
5. Ram air.
6. Inverters of 1,500-va (volt-ampere) capacity and above.

Also considered in the over-all characteristic is the type of drive with which par-

LIMIT 1		LIMIT 2		LIMIT 3		LIMIT 4		LIMIT 5		LIMIT 6	
SEC.	FREQ.	SEC.	FREQ.	SEC.	FREQ.	SEC.	FREQ.	SEC.	FREQ.	SEC.	FREQ.
0.00	380.0	0.00	380.0	0.00	380.0	0.00	420.0	0.00	420.0	0.00	420.0
0.01	375.0	0.01	377.5	0.01	378.75	0.01	421.25	0.01	422.5	0.01	425.0
0.05	355.0	0.05	367.5	0.05	373.75	0.05	426.25	0.05	432.5	0.05	445.0
0.2	320.0	0.2	350.0	0.2	365.0	0.2	435.0	0.2	450.0	0.2	480.0
1.0	348.0	1.0	368.0	1.0	376.0	1.0	424.0	0.92	432.0	1.0	452.0
5.0	380.0	3.0	380.0	2.0	380.0	2.0	420.0	3.00	380.0	5.0	420.0

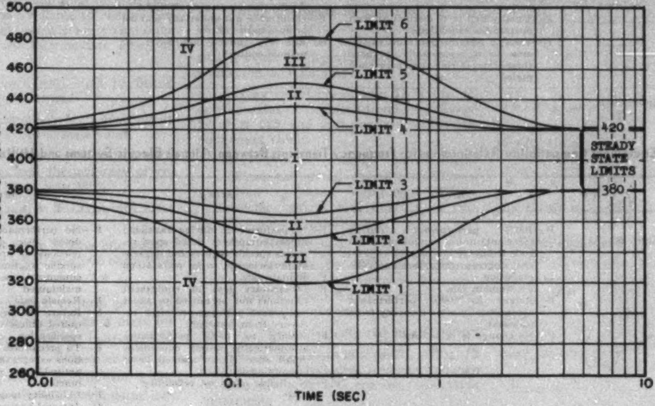


Fig. 12. Proposed transient frequency limits

alleling can be accomplished by frequency droop methods. Thus, the curves of Fig. 12 are intended to represent a composite frequency transient characteristic for all aircraft electric constant-frequency types of power systems. Compatibility of the limits in Fig. 12, between the electric system and utilization equipment, as considered in each of these four areas of electric-system operation, is abstracted in Tables I and II.

For proper utilization equipment design, knowledge of the rate of frequency change during frequency transients is required. In many cases, utilization equipment performance can be more influenced by small transient-frequency excursions at high rates of change, than by large transient-frequency excursions at low rates of change.

The available frequency transients data were reviewed for the purpose of extracting rates of frequency change. The rates of change during periods of excessive voltage change could not be considered accurate in the light of the instrumentation voltage sensitivity, and the time of response as previously discussed. The data did indicate that for frequency transients resulting from usual electric-system operations or area I, rates of change did not exceed 125 cps/sec; for frequency transients resulting from the occasional and more severe electric-system operations or area II, rates of change did not exceed 250 cps/sec; and for frequency transients resulting from the rare and extreme electric system operation or area III, rates of change did not exceed 500 cps/sec.

It is possible, and also likely, that frequency transients in area I have rates of change that are greater than 125 cps. These transients should be considered as being a transient associated with an area corresponding to its rate of change, rather than to its excursion. The definitions developed should consider that, whichever aspect of the frequency transient is most severe, the rate of change or excursion should become the governing aspect of classifying its area characteristic.

Frequency Modulation

Frequency modulation in aircraft constant-frequency electric systems originates from speed variations in a-c generator rotors due to shaft torsionals and drive speed regulation dynamics to maintain constant frequency. As long as the frequency modulation remains below reasonable levels in terms of its amplitude and rates, utilization equipment will not be affected by this characteristic. Experience has as yet not made completely clear as to what would be considered reasonable levels. Four-hundred-cycle inverters have been used for many years having considerable amplitude of frequency modulation but at unknown rates and with few recorded problems of incompatibility.

Tests were run for two inverters and hydraulic constant speed drives. These are adjusted in speed and load until the worst conditions of frequency modulation were obtained and recordings taken. Figs. 9 and 10 show the oscillograms obtained from the inverters. Fig. 11 shows the oscillogram obtained from the hydraulic constant speed drive. The instrumentation described in Fig. 8 was used for the upper trace and the instrumentation described in Fig. 3 was used for the lower trace of these recordings.

Field experience with incompatibility problems due to frequency modulation, although limited, have indicated that problems arise when the frequency modulation has gone to amplitudes above ±5 cycles. Very little data are available which verify that compatibility problems exist with frequency-modulation amplitudes below ±5 cycles.

Recent development of rate gyros operated from the aircraft electric power

Table I. Electric System Compatibility Relationships for Frequency Transients Between Aircraft Electric Systems and Utilization Equipments

Area 1	Area 2	Area 3	Area 4
D—Normal load switching F—All flight conditions F—Usual electric system operations required for function of aircraft and aircraft mission F—Occurrence under manual or automatic control	D—Bus switching D—Synchronizing D—Warmup F—Electric system operations required for function of electric system F—Usually tied in to distinct aircraft operational conditions F—Electric system operations in this area occur occasionally, perhaps three times during one operational period	D—Faults and recovery from faults D—Reset after fault not seen by utilization equipment F—Electric system operation in this area occurs rarely, perhaps once during any given aircraft flight F—Occurrence not under any manual or automatic control F—Recovery under automatic or manual control	F—No electric system operation in this area

Note: D = design, F = functional.

Table II. Utilization Equipment Compatibility Relationships for Frequency Transients Between Aircraft Electric Systems and Utilization Equipments

Area 1	Area 2	Area 3	Area 4
F—100% performance F—No change in reliability	F—100% performance (during transient) unless equipment detail spec defines allowance of regions and degrees of performance degradation F—Remain safe F—Return to 100% performance automatically after recovery from transient F—No change in reliability	F—No performance (during transient) unless equipment detail spec requires specific regions and degrees of performance to be maintained F—Remain safe F—Momentary loss of equipment functions not permitted to affect equipment performance after recovery from transient F—Return to 100% performance automatically unless equipment detail spec allows manual reset of equipment F—Negligible effect on reliability	F—No performance required unless detail spec dictates applicable portions of area 4 and requires specific regions and degrees of equipment performance to be maintained F—Remain safe F—Return to performance not required unless called out in detail specification F—To precipitate no adverse conditions except those caused by the natural loss of the equipment function F—Reliability may be affected D—Integral protection from overload

Note: D = design, F = functional.

system has pointed toward an increased requirement to contain limits of frequency variations to the closest practicable limits. This consideration is most applicable to frequency modulation since the modulation represents frequency variations which, during steady-state operation, occur continuously at rates of change critical to many equipments. Not only the amplitude but also the frequency (rate) characteristic of frequency modulation must be defined. A means of containing these aspects could be the use of a maximum limit of ±2 cps about a mean frequency for a given period of time to define the modulation amplitude. The ±2-cps deviation proposed as maximum excursion limits for frequency modulation leads to the consideration that maximum rate of change permitted for any complex wave of frequency modulation should be no worse than the maximum rate of change during a sine wave excursion of frequency modulation to the maximum allowed limits. If a continuous sine wave of frequency deviation with a one-second period is considered traveling to the extremes of ±2 cps about a nominal frequency, the rates of frequency change can be averaged as a slope going from one positive extreme of the frequency deviation to the adjacent negative extreme. This slope has an ordinate of 4 cps and an abscissa of 0.5 second, giving a rate 8 cps/sec. As shown by the derivations in Appendix II, there is a ratio between the slopes of a sine wave and triangular wave of 1.571 to 1 at the zero axis. Maximum slope of the sine wave, which occurs at its zero axis, is then approximately equal to 12.6 cps/sec. Thus, a maximum rate of frequency change should be set correspondingly at 13 cps/sec.

The frequency characteristic of frequency modulation is generally very random in its nature until one factor predominates, such as instability in a speed regulator of a constant speed drive. However, under usual conditions, the randomness of this characteristic does not allow for ready analysis of its sine wave components. Specifying the maximum rate of frequency change limits sufficiently the frequency characteristic of the frequency modulation. The maximum rate of frequency change can be easily identified as a maximum slope of a recording of frequency deviation and used as a means to determine conformance.

When a phase discriminator, such as shown in Fig. 8, is used as a means to record and analyze frequency modulation, correspondence must be obtained toward equivalent quantities obtained on a frequency basis. The actual shift in degrees differentiated in respect to time $d\phi/dt'$ is the function of frequency difference between the measured frequency and the standard frequency. The sense or direction of the frequency difference from the standard frequency has been lost. A phase shift slope of 720 degrees per second corresponds to a frequency deviation of 2 cps from the standard frequency. Rates of frequency change of the measured frequency corresponds to the change from one phase shift slope to another and as such is extremely awkward to handle. Because of this difficulty of determining frequency data from phase shift data, it is considered that frequency rather than phase instrumentation will be most applicable for the measurement of frequency modulation in constant-frequency electric system.

Proposed Instrumentation Criteria

It is proposed that instrumentation be developed which can be considered adequate for measuring frequency transients in an aircraft constant-speed electric system. Such instrumentation should be specified tentatively to provide the following over-all characteristics:

1. Output (voltage analog of input fre-

quency) to be independent of input voltage variations between the limits of 5 to 250 rms volts. The output with a constant-frequency input, and while voltage limits are varied, shall not vary more than 5% of full-scale value.

2. Output shall have center zero corresponding to the nominal input frequency of 400 cps with a maximum span of ±80 cps. Facility should be available to change the ±80-cps span to at least a ±4-cps maximum span and with an intermediate span of ±20 cps.

3. Any interference level shall be below 1% of any full-scale output.

4. Accuracy of any steady-state final output shall be within 2% of full-scale value.

5. For an input frequency deviation with a rate of 500 cps/sec, the accuracy of reproduction shall be at least 95%.

6. Output shall not be less than ±5 volts for full span.

7. Internal impedance looking into output terminals shall be less than 5,000 ohms with critical damping.

The instrumentation criteria proposed is suggested to be adopted as a basis for determining adequacy of instrumentation in measurements of frequency transients and frequency modulation. The criteria, when met by all instrumentation, could place all data on a sufficiently common basis to allow correlation.

Conclusions

It is concluded that the following limits and requirements be established for aircraft 380- to 420-cps electric systems:

CHARACTERISTICS OF FREQUENCY TRANSIENTS

1. Frequency transients as a result of usual electric system operations required directly for function of aircraft and aircraft mission shall be within area I of Fig. 12 and have rates of frequency change not exceeding 125 cps/sec.

2. Frequency transients as a result of occasional electric system operations required expressly for function of the electric system shall be within area II of Fig. 12 or less and have rates of frequency change not exceeding 250 cps/sec.

3. Frequency transients, caused by rare electric system operations whose occurrence is not under any manual or automatic control, shall be within area III of Fig. 12 or less, and have rates of frequency change not exceeding 500 cps/sec.

4. No electric system operation shall occur which will result in any frequency transients going into area IV of Fig. 12, or having rates of frequency change greater than 500 cps/sec.

CHARACTERISTICS OF FREQUENCY MODULATION

1. Frequency modulation shall be contained during any 1-minute period to be within a band of ±2 cps about a mean frequency. This band shall always be inside the steady-state frequency limits.

2. Rates of frequency change due to frequency modulation shall not be greater than 13 cps/sec.

REQUIREMENTS FOR UTILIZATION EQUIPMENT

1. During transients in area I of Fig. 12, or with rates of transient-frequency change less than 125 cps/sec, utilization equipment shall provide 100% performance and remain unaffected in reliability.

2. During transients in area II of Fig. 12 or with rates of transient frequency change between 125 and 250 cps/sec, utilization equipment shall:

   (*a*). Provide 100% performance, unless the detail specification for a given utilization equipment defines specific regions and degrees of performance degradation.

   (*b*). Remain safe.

   (*c*). Recover to 100% performance automatically after recovery from the transient when any degradated performance is allowed.

   (*d*). Remain unaffected in reliability.

3. During transients in area III of Fig. 12 or with rates of transient frequency change between 250 and 500 cps/sec, utilization equipment:

   (*a*). Shall have no performance requirements during the transient, unless the equipment detail specification requires the specific regions and degrees of performance to be maintained.

   (*b*). Shall remain safe.

   (*c*). May have momentary loss of function during the transient; however, this momentary loss shall not affect equipment performance after recovery from the transient.

   (*d*). Shall return to 100% performance automatically unless equipment detail specification permits manual reset of equipment after transient recovery.

   (*e*). Shall have negligible effect on reliability.

4. During transients in area IV of Fig. 12 or with rates of transient frequency change above 500 cps/sec, utilization equipment:

   (*a*). Shall not be required to perform.

   (*b*). Shall remain safe, during and after the transient.

   (*c*). Shall precipitate no adverse conditions except those caused by the neutral loss of the equipment function.

   (*d*). May be affected in reliability.

Appendix I

Derivations to determine error of fundamental component of a variable voltage sine wave clipped at a constant amplitude are given in the following.

Given:

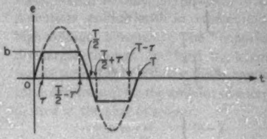
E_m = peak voltage of a variable-voltage sine wave

b = peak voltage at the constant clipping level

e = variable voltage at any instant

T = total period of a sine wave

τ = time at which clipping starts



$$e=E_m \sin \omega t,\ \omega=\frac{2\pi}{T},\ \frac{1}{\omega}=\frac{T}{2\pi} \qquad (1)$$

$$b=E_m \sin \omega\tau \qquad (2)$$

also

$$b=E_m \sin \omega\left(\frac{T}{2}-\tau\right)$$

From equation 2

$$\sin \omega\tau=\frac{b}{E_m}$$

and

$$\omega\tau=\text{arc sin}\left(\frac{b}{E_m}\right)$$

therefore

$$\tau=\frac{T}{2\pi}\ \text{arc sin}\left(\frac{b}{E_m}\right) \qquad (3)$$

General Fourier analysis of a periodic function is given by the expression

$$e(t)=B_0+B_1 \cos \omega t+B_2 \cos 2\omega t+\ldots+ A_1 \sin \omega t+A_2 \sin 2\omega t+\ldots \qquad (4)$$

In this case B_0=zero, since there is no d-c component. Also since in this case there are no even harmonics, $B_1=O$, $B_2=O$ ....

Hence in this case $e(t)=A_1 \sin \omega t+A_2 \sin 2\omega t+\ldots A_n \sin n\omega t$ where

$$A_n=\frac{2}{T}\int_0^T e(t) \sin (n\omega t)dt$$

To obtain the general expression for the Fourier coefficients of the discontinuous wave, the period O to T is divided into the intervals O to τ, τ to $T/2-\tau$, $T/2-\tau$ to $T/2+\tau$, $T/2+\tau$ to $T-\tau$ and $T-\tau$ to T.

$$A_n=\frac{2}{T}\left(\int_0^\tau E_m \sin \omega t(\sin n\omega t)dt+ \int_\tau^{\frac{T}{2}-\tau} b \sin (n\omega t)dt+ \int_{\frac{T}{2}-\tau}^{\frac{T}{2}+\tau} E_m \sin \omega t(\sin n\omega t)dt+ \int_{\frac{T}{2}+\tau}^{T-\tau} (-b) \sin (n\omega t)dt+ \int_{T-\tau}^{T} E_m \sin \omega t(\sin n\omega t)dt\right) \qquad (5)$$

Since we are interested only in the fundamental component after clipping, we make $n=1$ to derive the fundamental component. Then

$$A_1=\frac{2}{T}\left(E_m\int_0^{\tau}\sin^2\omega t dt+\right.$$
$$b\int_{\tau}^{\frac{T}{2}-\tau}\sin\omega t\,dt+$$
$$E_m\int_{\frac{T}{2}-\tau}^{\frac{T}{2}+\tau}\sin^2\omega t\,dt+$$
$$(-b)\int_{\frac{T}{2}-\tau}^{T-\tau}\sin\omega t\,dt+$$
$$\left.E_m\int_{T-\tau}^{T}\sin^2\omega t\,dt\right) \quad (6)$$

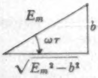
Integrating the expression, it becomes

$$A_1=\frac{2}{T}\left[\left(\frac{E_m\tau}{2}-\frac{E_m\sin 2\omega\tau}{4\omega}\right)+\right.$$
$$\frac{2b\cos\omega\tau}{\omega}+\left(E_m\tau-\frac{E_m\sin 2\omega\tau}{2\omega}\right)+$$
$$\left.\frac{2b\cos\omega\tau}{\omega}+\left(\frac{E_m\tau}{2}-\frac{E_m\sin 2\omega\tau}{4\omega}\right)\right] \quad (7)$$

Simplifying and collecting terms

$$A_1=\frac{2}{T}\left(2E_m\tau-\frac{E_m\sin 2\omega\tau}{\omega}+\frac{4b\cos\omega\tau}{\omega}\right) \quad (8)$$

From equation 2 one obtains



$$\cos\omega\tau=\frac{\sqrt{E_m{}^2-b^2}}{E_m} \quad (9)$$

Sin $2\omega\tau=2\sin\omega\tau\cos\omega\tau$; substituting from equations 9 and 3 obtains

$$\sin 2\omega\tau=2\left(\frac{b}{E_m}\right)\left(\frac{\sqrt{E_m{}^2-b^2}}{E_m}\right)$$
$$=\frac{2b\sqrt{E_m{}^2-b^2}}{E_m{}^2} \quad (10)$$

Substituting from equations 1, 3, 9, and 10 into equation 8 to remove the time function, the following can be derived

$$A_1=\frac{2}{T}\left[2E_m\frac{T}{2\pi}\arcsin\left(\frac{b}{E_m}-E_m\left(\frac{T}{2\pi}\right)\times\right.\right.$$
$$\left.\frac{2b\sqrt{E_m{}^2-b^2}}{E_m{}^2}+4b\left(\frac{T}{2\pi}\right)\times\frac{\sqrt{E_m{}^2-b^2}}{E_m}\right. \quad (11)$$

In simplifying, a final expression for the amplitude of the fundamental frequency is obtained

$$A_1=\frac{2E_m}{\pi}\arcsin\left(\frac{b}{E_m}\right)+\frac{2b\sqrt{E_m{}^2-b^2}}{\pi E_m} \quad (12)$$

For the condition when clipping starts, b equals E_m and the expression reduces to the term

$$A_1=\frac{2E_m}{\pi}\arcsin\left(\frac{E_m}{E_m}\right)+\frac{2E_m\sqrt{E_m{}^2-E_m{}^2}}{E_m}$$

and

$$A_1=E_m \quad (13)$$

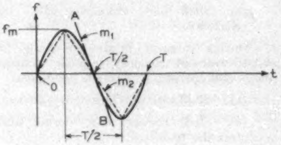
For the condition where E_m is infinitely large in relation to b (the clipping level), the expression reduces to $A_1=1.27\ b$. Thus, when the starting point of E_m is very large in reference to the level at which it is clipped, as it is reduced to the clipping level b there is a decrease of the fundamental component from 1.27 to 1. This constitutes an error as high as 0.27/1.27 or 21.3% in the assumption that squaring a wave of varying amplitude maintains a constant output at the fundamental frequency.

Appendix II

Derivation to determine the ratio to slope of a triangular wave to the slope of a sine wave at its point of inflection, both waves having the same time period and maximum amplitude is given in the following.

Given:

f_m = maximum deviation in terms of frequency
f = variable frequency at any point
T = total time period of both the sine wave and the triangular wave
m_1 = slope of sine wave at the point of inflection
m_2 = slope of triangular wave



$$f=f_m\sin\omega t,\text{ where }\omega=2\pi f\text{ and }\omega=\frac{2\pi}{T} \quad (14)$$

Differentiating the equation 14 with respect to t, one obtains

$$\frac{df}{dt}=\omega f_m\cos\omega t \quad (15)$$

Differentiating equation 15 with respect to t, one obtains

$$\frac{d^2f}{dt^2}=-\omega^2 f_m\sin\omega t \quad (16)$$

For the point of inflection along the t axis

$$\frac{d^2f}{dt^2}=0,\ \therefore\omega^2 f_m\sin\omega t=0 \quad (17)$$

From this equation it may be shown that the point of inflection and maximum slope along the t axis is $T/2$.

To determine the slope at the point of inflection, substitute $t=T/2$ in equation 15 obtaining

$$\frac{df}{dt}=-\frac{2\pi}{T}f_m=m_1 \quad (18)$$

From the figure it is evident that the slope of the triangular wave going through the point $T/2$ is given by

$$m_2=-\frac{2f_m}{T/2}=-\frac{4f_m}{T} \quad (19)$$

The ratio of slopes is then as given in the following

$$\frac{m_1}{m_2}=\frac{\pi}{2}=1.571 \quad (20)$$

References

1. FREQUENCY MODULATION AND LOAD-DIVISION INSTABILITY IN 400-CYCLE AIRCRAFT ELECTRIC SYSTEMS, Henry Oman. *AIEE Transactions*, vol. 75, pt. II, July 1956, pp. 181–88.

2. CHARACTERISTICS OF AIRCRAFT ELECTRIC POWER. *Military Specification MIL-E-7894A (ASG)*, May 17, 1955.

3. PHASE-SELECTIVE DETECTORS, Curtis R. Schafer. *Electronics*, New York, N. Y., Feb. 1954, p. 188.

4. THEORY AND APPLICATION OF INDUSTRIAL ELECTRONICS (book), John M. Cage, C. J. Baske. McGraw-Hill Book Company, Inc., New York, N. Y., 1951, p. 223.

5. LABORATORY PHASE COMPARATOR, D. Feinstein. *Electronic Design*, New York, N. Y., July 1954, p. 30.

Electric Ignition for Conductive Fuels

S. J. SHEHEEN
ASSOCIATE MEMBER AIEE

Synopsis: The design and operational environment of the engines in today's jets, rockets, and missiles have raised new problems in the field of ignition. In the jet engine, the ignition must function at both high and low altitudes. At greater altitudes higher energies are required to offset decrease in oxygen: Higher energies naturally mean larger and heavier units, but size and weight are critical design factors which are strictly limited by engine specifications. In rocket and missile engines the problem is rendered more acute by the use of nitric acid or other agents which make the fuel electrically conductive. Conventional high-voltage ignition systems with jump-gap plugs are not adaptable to these fuels.

TO DEMONSTRATE the effects of fuel conductivity it is only necessary to point out the effect of conductive moisture on an ordinary automotive jump-gap plug Any attempt to provide a spark across the gap by a high-impedance source such as a coil is known to be futile. The moisture itself forms a shunt between the electrodes, which prevents the build-up of the necessary breakdown voltage. With reference to the gap together with the leakage path as the load, it can be stated that the amount of energy transferred from the source is a function of the load. In this case, because of the conductive shunt, it is low; the moisture destroys the over-all efficiency of the system. Aside from this simple aspect of the problem, the ignition system must cope with the condition that exists when the conductivity of the fuel at the point of ignition is not constant.

In order for a rocket to obtain the tremendous acceleration necessary for flight, the final fuel pressure may attain hundreds of pounds. The fuel enters the ignition chamber in the form of a cone whose angle varies over wide limits. This angle, under normal atmospheric conditions, depends on the fuel pressure. The surface of the cone is usually a solid stream of fuel while the interior is a turbulent mixture. This means that during the period when ignition must take place, the conductivity across the gap may vary from a few ohms to megohms.

At the same time it is essential to provide positive ignition in a very brief period of time, otherwise an excessive build-up of fuel may occur and cause an explosion elsewhere than in the combustion chamber.

Two possible approaches are available for igniting conductive fuels; these are pyrotechnical devices and thermoelectric or electric discharge systems. Of the two types, the latter is preferable because it provides the necessary fast and repetitious source of energy.

The Design of the Source

Under such conditions, the design of the ignition system must be such as to transfer the maximum amount of energy to the electrodes regardless of whether they are situated in a highly atomized vapor or in a liquid stream of conductive mixture.

Since the total amount of energy available is used partly in the source impedance and partly in the load, two methods of meeting the variable conductivity factor can be considered: varying the source impedance to balance changing load conditions; and providing a low source impedance and introducing an auxiliary shunt to compensate for variations in the load. It is not impossible to meet the requirements of the first method; in fact, some jet ignition systems utilize this design. The source impedance actually changes from a high to a low impedance to meet variations in load resistance and to provide sufficient energy for an arc discharge. The design, however, then becomes more elaborate, both as to the source and to the igniter required.

The second alternative is more attractive for the following reasons:

1. It can provide enough current to arrive at a high-energy arc discharge.
2. If the impedance of the source is low, the energy transferred to the igniter which is shunted by the conductive fuel is high.
3. A shunt gap igniter can be designed to provide a consistent low-impedance load, regardless of the conditions of the fuel.

The capacitor discharge system, Fig. 1, satisfies all these requirements. It is simple in design, containing vibrator or motor-driven breakers, transformer, rectifier, and a capacitor which acts as a source of low impedance. The capacitor is normally charged to 1,000 or 2,000 volts, which is only 10 per cent of the potentials encountered in conventional high-voltage systems. The capacitor discharges through the load by a contact or series gap. The size of the capacitor is determined by the amount of energy desired at the plug or igniter.

The Igniter

The low system voltage is made possible by the design and performance of the shunt igniter used with it, which is illustrated in Fig. 2. The conductive path between the center and ground electrodes consists of a semiconductor material, which replaces the insulating alumina body customary in the ordinary jump-gap plug. At atmospheric pressure, this path breaks down and provides a spark at approximately 500 volts, as compared with 10 to 15 kv required for a jump-gap of the same dimension. Operation at such low voltage has the obvious advantage of reducing insulation problems throughout the system. This shunt igniter also offsets numerous fouling conditions which render a jump-gap plug ineffective in a jet engine, as it is but slightly affected by wide variations in pressures, carbon deposition, and wax formation at very low temperatures.[1]

This igniter was originated in 1950 by Wijtze Beye Smits of Voorburg, Holland. He deliberately utilized the carbon deposits on the surface of an insulator for electrical leakage to promote breakdown. The igniter was designed with a gap of less than 0.010 inch. This method necessitated new deposition after each successive or series of sparks until a "carbon track" or permanent deformation was formed in the body of the insulator by the discharges. This development lead to extensive research in the igniter field by ignition manufacturers in an attempt to minimize the problems presented by the turbojet.[1]

Because of the excessive high temperatures and variations in the atmosphere of the jet burner, some ignition manufacturers retained the original idea of the

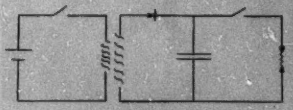


Fig. 1. Basic capacitor discharge system with shunt gap

Paper 57-492, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, March 7-9, 1957. Manuscript submitted August 29, 1955; made available for printing March 15, 1957.

S. J. Sheheen resides in Ithaca, N. Y.

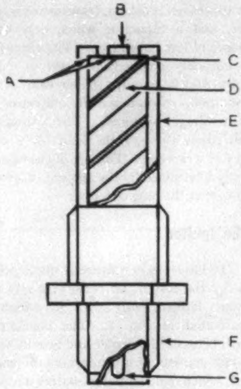


Fig. 2. Shunt igniter

A—Metal semiconductor contact areas
B—Center electrode
C—Semiconductor
D—Insulator
E—Shell
F—Adaptor
G—Center contact pin

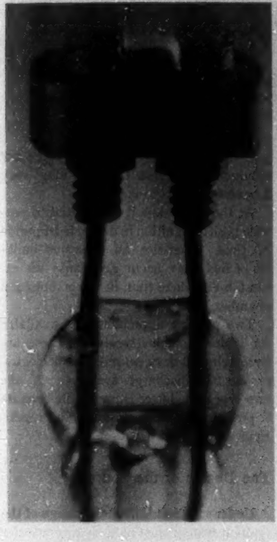


Fig. 4. Initial emission at contact prior to complete breakdown of the gap

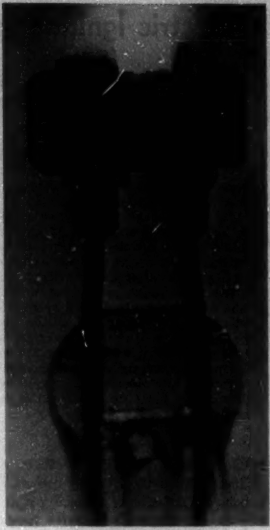


Fig. 5. Arc discharge across the gap

carbon-tracked small-gap-surface igniter, while others depended on refractory oxides to obtain a positive semiconducting surface as the initial path for the energy from the source. This surface was fired either as an integral part with the green insulator body or as a coating on or impregnation in a baked ceramic.

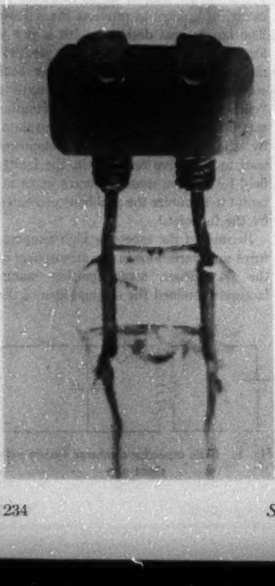


Fig. 3 (left). Symmetrical shunt gap

The low-resistance shunt path is provided by the oxygen difficiencies or excess for the cations of the refracting oxides, which have more than one valance state. another method of obtaining conductivity in an otherwise insulating refractory oxide is to substitute in the lattice a secondary cation of an oxide whose ionic radius and structure will allow it to replace a cation of the primary oxide; the valance of the secondary being greater or less than that of the primary metal.[3]

The results of these investigations have led, as far as terminology is concerned, to two types of igniters: shunt-gap and surface-gap igniters. The shunt-gap igniter has a much lower breakdown voltage, usually in the region of 300 to 1,000 volts. The normal gap lengths range from 0.030 to 0.100 inch. The disadvantage is in the semiconducting body. Since the formation of the shunt between the two electrodes for the igniter depends on the deficiency or excess of oxygen in an oxide, continued heating at high temperature of an excess oxide (*P*-type) in a reducing atmosphere may lead to higher resistance and onset voltage, and similarly, a high-temperature oxidizing atmosphere in the vicinity of an oxygen-difficient oxide (*N*-type) for a prolonged period of time will cause the onset voltage to increase.

The normal firing or baking temperature of some refractory oxide surfaces ranges from 2,500 degrees to 3,000 degrees Fahrenheit, which is of the same order as for sparkplug insulators. The arc temperature during breakdown above the semiconductor exceeds this, but the duration is very short. Very little thermal energy is transferred to the surface from the arc as compared to that developed at the metal contact, where the greatest amount of erosion and decomposition takes place.

The surface-gap igniter operates in the same manner as the shunt igniter. Their difference lies in the method of obtaining conductivity for the surface. To obtain conductivity for this igniter, a pencil is deliberately applyed on the surface of the insulator separating and in contact with the two electrodes. This promotes conductivity and allows the initial surface currents influenced by a charged capacitor to track between the electrodes. Contined effort and sufficient voltage will finally change the crystal structure of the insulators so as to embody minute semiconductive paths. The conductivity of

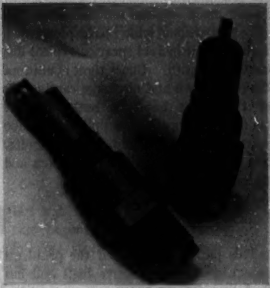


Fig. 6. Shunt igniter with auxiliary electrode

these tracks is much lower than that of the semiconductor of the shunt igniter, and thus limits the size of the gap and increases the onset voltage. The onset voltage for the normal 0.010-inch gap of the surface igniter ranges from 1,500 to 3,000 volts.

A basic theory behind the breakdown of the gap in a shunt igniter is quite different from the action of a conventional jump-gap plug, as will be evident from the comparison which follows. In the jump-gap plug, breakdown of the gap is initiated by free electrons increasing their kinetic energy at the expense of the applied potential between the electrodes. In doing so the electrons collide with other atoms and molecules and expend part of their energy in ionizing them. This is a cumulative process as long as potential is applied. The greater the number of electrons, the greater the possibility of producing more free electrons. The increase, which is exponential, continues until complete ionization occurs, causing a breakdown of the gap.[3] In this plug a large gradient is necessary to provide enough free electrons so that ionization and eventual breakdown may occur.

The theory of the means by which the shunt igniter provides ionization is quite different. When a semiconductor, for example, an N-type, is in contact with a metal, a potential exists across the interface, and it is approximately equal to the difference of the work functions of the materials. This contact potential exists because the electrons in the conductive band of the semiconductor migrate to a lower potential in the metal, the work function of the semiconductor being smaller than the metal. Thus there exists an interchange of electrons across the contact at equilibrium, but more electrons go from the semiconductor to the metal than from the metal to the semiconductor. The electrons of the metal have to gain energy to jump across the contact. The result of this exchange leaves a positive charge on the semiconductor and a negative charge on the metal at the contact surface. This is the contact potential and its value is dependent on the difference of the work functions.[2]

Fig. 7. Breakdown of a conductive liquid at auxiliary electrode during immersion in nitric acid

As the result of this contact potential, a unilateral impedance is formed and current can flow with greater ease in one direction than in the other, depending on the direction of the applied voltage. If the applied voltage is such as to cause electrons to go up the potential hill, heat will develop at the contact surface. If enough electrons can be made available to go up this hill, heat can be increased until incandescence occurs at the interface and a multitude of electrons will be emitted. These initial electrons are emitted in the electrical field between the center and ground electrodes, where they increase their energy and produce more electrons until breakdown occurs. Thus by charging the capacitor to the value determined by the energy needed at the igniter and discharging it through the shunt igniter, breakdown will take place.

The ability of a contact to produce sufficient initial emission at low voltage is one of the determining factors in its selection. A figure-of-merit for a semiconductor-to-metal contact would be the ratio of emission to the power dissipated per unit area.

Some of the experiments seem to support an existing theory that the efficiency of an N-type is greater than the P-type in the production of initial electrons provided the metal in contact, the number of carriers, and conductivity are the same.

It has been established that the potential existing at a semiconductor-to-metal contact is as follows:

For N-type

$$\varphi_N=\frac{W_M-W_N}{e} \tag{1}$$

and for P-type

$$\varphi_P=\frac{W_P-W_M}{e} \tag{2}$$

provided $W_M > W_N$ for 1 and $W_P > W_M$ for 2. Otherwise an ohmic impedance is presented without a barrier.[4] In equations 1 and 2:

W_M = work function of the metal
W_N = work function of the N-type semiconductor
W_P = work function of the P type semiconductor
e = charge of an electron
φ_N = contact potential at N-type contact
φ_P = contact potential at P-type contact

To complete the comparison, φ_N should be made equal to φ_P and the same potentials applied in the reverse direction. The current densities will be equal.[4]

$$J=\sigma\epsilon\left(^{\frac{-e\varphi}{kT}}\epsilon^{\frac{eV}{kT}}-1\right) \tag{3}$$

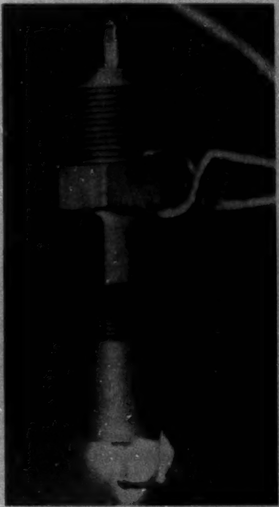


Fig. 8. Breakdown at shunt gap in a nitric acid spray

where

J = current density
φ = contact potential
k = Boltzmanns constant
T = temperature
V = applied voltage
σ = conductivity

From equations 1, 2, and 3, it can be seen that the power dissipated per unit area and the resulting temperatures will be approximately the same. The thermionic emission for the metal as the results of these temperatures at the barrier is

$$I = A_0 T^2 e^{\frac{-W_m}{kT}} \qquad (4)$$

The equation for emission from a semiconductor is of the same form:

$$I = A_1 T^2 e^{\frac{-W}{kT}} \qquad (5)$$

where W is the work function of the semiconductor.[2]

Since $W_P > W_N$ it can be concluded from equations 1-5 and the foregoing conditions that the ability of an N-type semiconductor to produce initial electrons is greater than that of the P-type and thus the N-type can have a lower breakdown voltage.

Using a symmetrical shunt gap as shown in Fig. 3, Figs. 4 and 5 are actual photographs of two stages during breakdown. The two contacts are of the same material and size, the gap length is 0.100 inch, and the energy available is 800 volts on a 4-microfarad capacitor, charged at atmospheric pressure. In Fig. 4 the initial emission of electrons is taking place at the interface on the right between the semiconductor and the metal. This action takes place during the early stage of breakdown, and emphasizes the condition of unilateral impedance imposed on the shunt gap. In Fig. 5 breakdown of the gap has resulted, and complete high-energy discharge is taking place.

Auxiliary Electrodes

This type of shunt igniter performs satisfactorily when the conductivity of the shunt is somewhat greater than that of the fuel, when the electrodes are located in a turbulent mixture or spray. When they are immersed in a solid stream of liquid, however, the fuel conductivity may be so high as to bleed off energy rapidly enough to prevent breakdown. To make certain of ignition under such conditions, two auxiliary electrodes are added to the shunt igniter, as illustrated in Fig. 6. The function of these electrodes is to confine a sufficient amount of the energy in a small space so that rapid electrolysis can occur.

As the fuel is electrolytic, application of a potential to the electrodes will not cause the gap to break down near the surface of the igniter, but will cause rapid decomposition of the electrolyte near the auxiliary electrodes. The large current provided by the low-impedance source, i.e., the capacitor, causes a hot spot to appear on one of the electrodes. This action is similar to that which occurs during anodic burning.

Even when fuel conditions permit the use of the simple shunt igniter, the addition of auxiliary electrodes has advantages. Operation in a conductive fuel produces great thermal shock and high pressures, which are diverted from the semiconductor gap to the electrodes.

Figs. 7 and 8 illustrate the functioning of a shunt igniter equipped with auxiliary electrodes under the two extreme conditions of a conductive media. In Fig. 7 the iginiter is completely immersed in nitric acid. The breakdown can be seen taking place between the electrodes. In Fig. 8 the same igniter is exposed to a turbulent nitric acid spray, a typical discontinuous path. Under these conditions the breakdown and discharge can be seen at the shunt gap. In both of these tests the energy was supplied by 1,400 volts on a 4-microfarad, and the igniter had the shunt gap set at 0.050 inch and the auxiliary gap at 0.030 inch.

Conclusions

This paper has outlined the problems encountered in conductive fuels and has demonstrated a design of an ignition system consisting of a low-impedance source and a shunt-gap plug with auxiliary electrodes adapted to the ignition of fuels having high or variable conductivity. The shunt and surface igniters, with their high-energy systems, are being used to a great extent in jet aircraft of the United States, England, and Canada.

The repetition rate and energy per spark from the low-impedance source necessary for operation in a reciprocating engine, and other disadvantages of the shunt igniter, prevents the high-energy system from giving the conventional low-cost, high-voltage system any serious competition.

The combination igniter consisting of a shunt and auxiliary electrodes has been successful in experiments with conductive fuels and is being investigated further.

References

1. HIGH-ENERGY IGNITION SYSTEMS FOR GAS-TURBINE STARTING. Avon, Rotax, Ltd.

2. SEMI-CONDUCTORS (book), D. A. Wright, Methuen and Company, Ltd., London, England, 1950.

3. GASEOUS CONDUCTORS, J. D. Cobine.

4. CRYSTAL RECTIFIERS, H. C. Torrey, C. A. Whitmer. *MIT Radiation Series 15*, McGraw-Hill Book Company, Inc., New York, N. Y.

Ground-Power Equipment for the Support of Air Force Weapons

T. B. HOLLIDAY
FELLOW AIEE

THE MATERIAL that is presented in this paper was gathered during an engineering study on the same subject which was made for the Air Research and Development Command of the United States Air Force. The objective of the study was preparation of recommendations for needed development in the area of ground support for secondary power. In aviation terminology, secondary power is defined as that power which is used for purposes other than the propulsion of aircraft and missiles. Aircraft and missiles are defined as weapons, and a weapon system includes everything that is needed for maintenance, calibration, testing, and training during operational use of the weapon. "Ground power" substitutes for air-borne secondary power during tests. The familiar secondary power systems are electric, hydraulic, and pneumatic. During the study, it was found desirable to treat air conditioning as a fourth and newly important type of secondary power.

Comments were obtained from all combat and technical Commands of the Air Force, from all weapon manufacturers who serve the Air Force, from engine manufacturers, from manufacturers of secondary power-generating equipment, and from some manufacturers of equipment that uses secondary power. In all, more than a half million items of information were gathered. The summary and collation of these data revealed a pattern of future needs and suggested desirable solutions to those needs.

In order to make a prediction of the probable needs in ground power, it is first necessary to predict the trends of the air-borne secondary power systems. These trends cannot be estimated without first making some assumptions, if not predictions, as to the trend of aircraft and missiles. This was the approach that was used during the Air Force study, and it produced much information of indirect value, which was to contribute to the planning of ground power. The results as they pertain to secondary power and the support thereof are presented here.

This article is limited to a discussion of ground electric power. It was found that weapons of the future can be served by five families of power units: (1) ramp unit, (2) trailer-mounted mobile unit, (3) stationary or portable unit, (4) fixed base power unit, and (5) electric motors. Of these, all but one are primarily sources of electric power. The second item will be found most useful as a hydraulic power unit. It became apparent that the turbine prime mover will rapidly supplement and then supersede the reciprocating engine. Thus, it was concluded that future developments should be based on the turbine drive.

There are two primary problems in any effort to plan ground power support. First, the effectiveness of the weapon must not be impaired. Second, and to save costs, as many common items as possible should be developed to conserve development time, personnel, and funds.

General Predictions

About the only safe prediction that can be made with regard to aviation development is that any given period will yield about twice the technological development which was accomplished during the same length of time in the immediate past. It is becoming increasingly apparent that there is no limit to the flight altitude and speed that will be achieved. The thermal barrier will be overcome by developments in materials, or by flight at higher altitudes. The developments in missiles will lead the way into outer space, but soon afterward, the missile will be converted to a manned aircraft of essentially similar performance.

Flight into outer space will advance the limits of environmental conditions. Higher flight speeds brings increased aerodynamic forces and the need for greater power in secondary power systems for the servo controls. Contributing to the achievement of higher speeds will be more compact designs of fuselage and wings. Yet more and more secondary powered equipment will be installed in the reduced space. Heat losses from that power must be carried away. The result indicates a much greater importance for air conditioning in the future. Development of components that can operate in higher ambient temperatures will reduce the power required for air conditioning, but the need will probably continue.

The higher speeds of the future will bring new problems in precision. Navigation, whether by human skill or by missile robot, will need greater precision. The necessary accuracy starts in the secondary power system, particularly in the electric system. It is evident that there will be a constant demand for more and more preciseness of control in the future. It is not likely that the aviation industry will ever be satisfied, because still more accuracy will always be sought in guidance equipment. Additional and supplementary precision is now built into the electronic subsystems. If more accuracy can be provided for the input secondary power system, less will have to be added by electronic means. Heretofore, only the electric system has been the subject of the demand for greater precision, but it is likely that other types of secondary power will tend toward greater precision in the future.

The hydraulic system will be the next to feel the pressure toward better control. Hydraulic power is the ideal means for servo control, as it provides ample power with light weight, is well suited to exerting lineal forces, is reasonably fast in response, and is not resilient or soft. Electric devices do not lend themselves to lineal actuation through any but very short distances, and most of them are too slow to be used as servo controls. The pneumatic system has too much resiliency to be used in continuous-duty servo applications, although it is very good for one-time applications.

The transistor will reach a major plateau of its usefulness by 1962. It has not brought all the blessings hoped for.

It will permit a great reduction in the size and power consumption of many electronic systems, particularly digital computers. This will reduce the amount of air conditioning that is needed to remove electric heat, but it will demand more precise temperature control. Many transistors have given characteristics over a small range of temperature. To achieve the ultimate precision in the electronic system, it will be necessary to control temperature around the transistors to quite narrow limits.

The power rating of the electric system

Paper 57-500, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, March 7–9, 1957. Manuscript submitted January 28, 1957; made available for printing April 8, 1957.

T. B. HOLLIDAY is with the Holliday Company, Elmhurst, Ill.

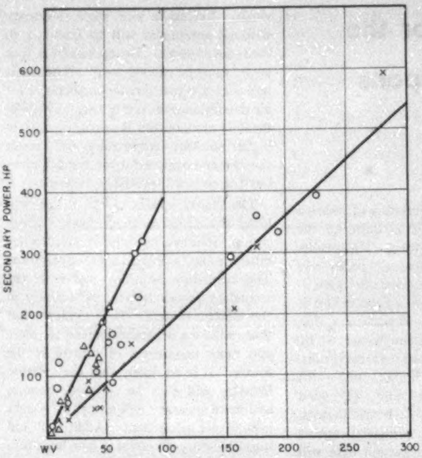


Fig. 1. Total secondary power load with respect to product of weapon gross weight and speed in Mach number

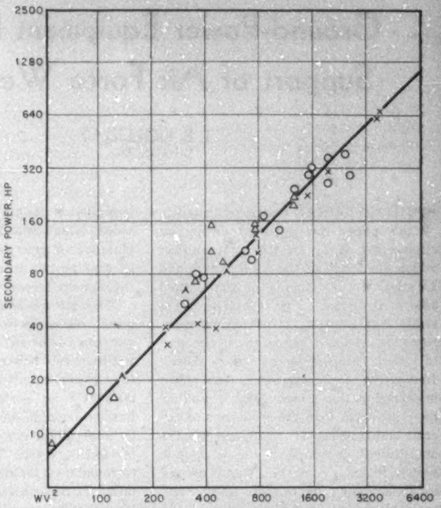


Fig. 2. Total secondary power load with respect to product of weapon gross weight and square of speed using indicated air speed

will continue to grow despite the advent of the transistor, but its advance in size will not be as dramatic as in the last 10 years, nor will it keep pace with the growth in weight and speed of aircraft. Relatively, the electric system will lose ground in quantity of power.

The hydraulic system will have a rapid growth in the next 10 years. Aircraft are now reaching speeds at which manual control is out of the question. Power boost must be used. Until now, it has been possible to use aerodynamic forces to counterbalance control forces, so that manual effort could control the flight. Power boost has been added to make the job easier. Now that aerodynamic forces are too large to use in this manner, and since power boost is always available, the aircraft designers may elect to design instability into the aircraft deliberately in order to attain still higher performance. This will increase the power rating of the secondary power system, and will force installation of a dual system in many types of aircraft, in order to have a safety factor for emergencies. A tenfold growth of hydraulic installed power in the next 10 years is predicted. Much of this will be due to the increased weight and speed of the aircraft, and some from increased uses for hydraulic power. Where three or four hydraulic subsystems have been used in the past, e.g., wheel brakes, landing gear retraction, flap actuation, and autopilot servos, as many as 12 can be expected in the next 10 years. Some of these will be continuous-duty applications in which hydraulic motors do work similar to that done by electric motors.

Low-pressure pneumatic power has been tried in the past 5 years and seems to be settling into a place of usefulness. It is tempting to use bleed air from the turbine engines through a pneumatic coupling to drive generators of secondary power. Although it is not an efficient coupling, it does remove accessories from the engine and permits a reduction in the frontal area of the engine. Advantage could also be taken of the ducts for deicing air as a path for pneumatic power. But, it is likely that deicing will be omitted on any aircraft that can cruise at altitudes beyond icing conditions. Todays's piston engine aircraft which are not able to evade all icing conditions, use deicing equipment for less than 3% of the total flight hours. Military aircraft need only pass through icing conditions enroute to or from iceless environment, and do not need deicing equipment at all. This means that the deicing ducting which could convey pneumatic power will not be installed. It is anticipated that the low-pressure air-turbine-driven devices will be restricted to emergency power generators and to missiles. The free turbine drive within missiles can use hot gases bled from the rocket engine, or ram air pressure.

High-pressure pneumatic power can be used in many special one-shot jobs, and as stored energy in the smaller missiles. This is a very economical way in which to carry stored power. The maximum pressure in such storage is likely to exceed 10,000 psi (pounds per square inch) in the next 10 years. The foregoing has given some general predictions as to air-borne secondary power systems. Some detailed characteristics are now considered.

The air-borne auxiliary power unit will have new importance in the future. Trends in weapon propulsion show that a static device may be the source of thrust in future designs, thereby removing the source of rotary power for electric generators. The rocket engine in missiles and the experimental aircraft is a forerunner of this trend. The increased use of the ram-jet engine is another. Space travel will bring entirely new operational concepts. A separate and independent source of electric power is and will be needed in such aerial vehicles.

Detailed Predictions

It was found that the total of secondary power has a relationship to the weight and speed of the weapon. In military aircraft, there is a small amount of power, as for radio communications, that may be said to be common to all aircraft without

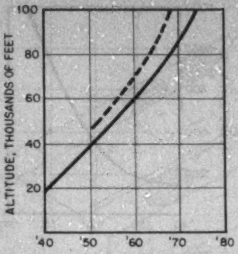


Fig. 3. Trend of operating altitude for manned aircraft (solid line) and air-breathing missiles

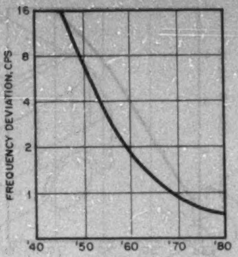


Fig. 5. Trend in reduction of frequency deviation (half of total tolerance)

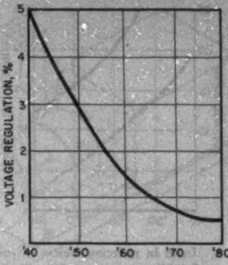


Fig. 6. Trend in reduction of voltage regulation (half of total tolerance)

regard to size. However, the amount of such power is almost negligible and is considered to be a part of the total for this estimate. The total power that is to be considered includes electric, hydraulic, pneumatic, and air conditioning. Any attempt to predict the amount of any one type of secondary power will result in confusion, because there are many accessory tasks that can use different types of power, and the various weapon manufacturers will choose different types of power for these. For example, the motor that drives an air-borne electric inverter can be electric or hydraulic.

It is necessary to consider only the power delivered by the secondary power generators. An evaluation that is based on the input power to these generators would be made erratic by the differing efficiencies of driving couplings. For example, two alternators might deliver 75 kw, which is an output power of approximately 100 hp (horsepower). A direct driven alternator might deliver this output with an input of 125 hp. The addition of a variable speed drive could demand an input power of approximately 140 hp, and the use of a pneumatic coupling to a driving turbine could require still higher input power. To avoid these several choices, only the output power of the secondary generators and pumps is considered.

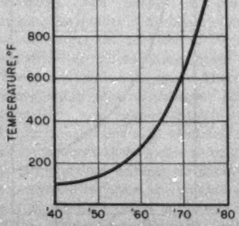


Fig. 4. Trend of environmental temperature around some items of air-borne electric equipment

The first attempt to make this analysis used secondary power as a function of the product of weapon weight and speed. The equation

$$P=kWV$$

where k is a constant, was used. The results of a plot of many aircraft is shown by Fig. 1. The ordinate of this plot is the total secondary power in horsepower. The abscissa is the product of gross weight in thousands of pounds and of velocity in Mach numbers. Two paths of correlation appear. Most of the supersonic aircraft lie on the upper line. It is obvious that this approach is not satisfactory, and that secondary power is dependent on an exponent of the velocity factor.

To bring better correlation, three steps were taken. First, indicated air speed was used instead of Mach number; second, the square of the velocity was used; and finally, some allowance was made for differing practices among weapon manufacturers with regard to hydraulic servo control systems. When dual systems were installed, the power for only one of them was counted. These changes brought the two curves together as shown by Fig. 2. This correlation is a reasonably straight line that can be projected into the future with some degree of certainty. This prediction of total power is important in the planning of ground power units, because the power rating of the servicing power unit will have some relation to the installed secondary power. After studying the highest single loads of the systems combined with the air-conditioning requirements that would accompany testing of those loads, it was decided that the ground power electric servicing unit would need to have about half the power rating shown for the total secondary system.

The air-conditioning loads of commercial transports cause their position on the plot to be high in the case of the older designs. As we move into the era of jet transports, the secondary power requirements come into line with those of military aircraft. This shows that the human factor becomes of less importance with higher performance in so far as air conditioning is concerned. In the future, the air-conditioning needs of the weapon itself will far overshadow those for the personnel in the weapon.

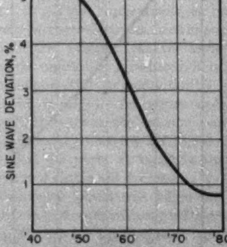


Fig. 7. Trend in reduction of factors that affect wave shape

The operating altitudes of weapons directly concern the environment for which secondary air-borne power equipment must be designed. These are of indirect interest in the ground power unit, for it will often be required to drive air-borne generator units in the servicing of the weapon. A projection of the trend in operating altitude of weapons is shown by Fig. 3, where the solid line represents aircraft, while the dotted line represents missiles. This plot shows an operational

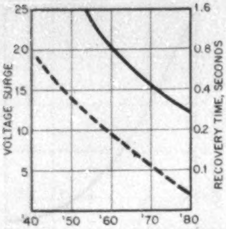


Fig. 8. Trend in reducing factors of voltage surges. Solid line is amount of surge in either direction and dotted line is time to recover to regulated values

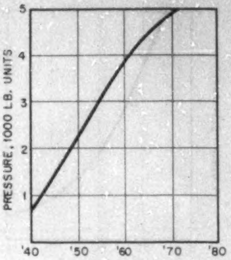


Fig. 10. Trend in pressure of air-borne hydraulic system

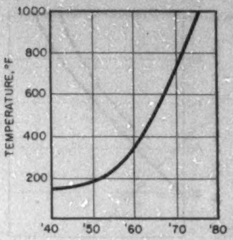


Fig. 11. Trend in hydraulic fluid temperature for normal operation

prediction, and not that of test and experimental flights and weapons. The experimental flight of today may be operational in another 10 years. It is already being required that air-borne secondary power generators be designed to be free of environment, and it is apparent that many more items of the secondary systems must follow suit in the next 20 years.

The environmental temperatures around critical pieces of equipment, particularly generators and pumps, will increase steadily, approximately as shown by the chart of Fig. 4. The temperature environment for most other items of secondary equipment will be less severe. Many of them will be located in space that is air conditioned. Others will be installed in space that is not close to heat generating sources of power.

Electrical Characteristics

The problems facing the designer of air-borne electric-power-generating equipment can be summed in one word, "precision." For basic power, the accepted standard is 3-phase alternating current at 208/120 volts and 400 cps (cycles per second). Generation of d-c power is rapidly fading out, although the need for direct-current is not. The d-c power that is needed will be delivered by transformer-rectifier units. There are some interesting possibilities for the direct generation of electric power, and they are likely to produce direct instead of alternating current. A technical breakthrough of this nature could radically alter air-borne electric systems.

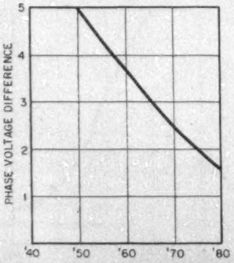


Fig. 9. Trend in reduction of differences between phase voltage of 3-phase system

Some consideration is being given to higher voltage systems that may be needed in larger aircraft of the future. The industry has tentatively agreed on a 3-phase 440-volt system for the proposed applications. There are several applications for other and higher frequencies, and they are brought about by the need for faster response from magnetic amplifiers. These frequencies have included 800, 1,200, 1,600, 2,000, 2,400, and 4,800 cps. If such special frequencies continue to be needed, it should not be a multiple of the basic power frequency, 400 cps. One frequency that should be avoided is 2,400 cps, for that is the ripple frequency of a 3-phase transformer rectifier unit. Successful development of power transistors that can replace the function of magnetic amplifiers will reduce or eliminate the need for special frequencies.

The preciseness of frequency control was one of the great debates encountered during the survey. There are four schools of thought. The first believes there is limited application on turboprop engines for variable-frequency alternators to supply raw or basic power. The second calls for a system basic power in which the frequency is controlled in a total range of 20 cps, 400 ±10 cps. The third group wants better control, and it is the prediction of this group that is shown by the plot of Fig. 5. The graph shows the permitted variation in each direction, half the total variation. All of these groups are fully aware of the need for very precise frequency control in critical electronic systems, but they believe that the refinement can and should be done in the power supply of the electronic equipment. The group contends that it is just as easy to refine all of the power at the outset, rather than a part of it.

Where the drive for precision in frequency will end is not easy to predict. Temperature-controlled crystals can deliver a precision in the order of 10^{-4}, which means of precision of 0.25×10^{-4} cps. The demand of the precision school today is for 0.01 cps, but there is little reason to doubt that the ultimate in precise control will be wanted eventually. The extreme precision is needed in missile guidance systems which cannot be countered. Such precision in a reference signal from the electric system can take the place of other references and is much easier to adapt to the guidance equipment.

Greater accuracy in the control of voltage will also be sought. The goal of weapon manufacturers is shown by the Fig. 6 graph. Such precision can be realized at only one accurate voltage control is not high, and is not a factor.

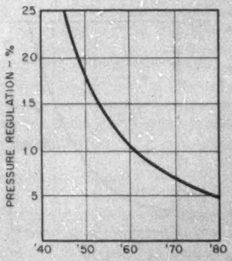


Fig. 12. Trend in pressure regulation (total tolerance)

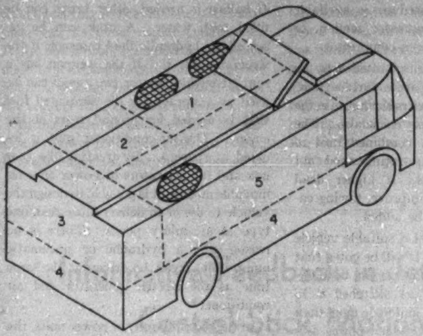


Fig. 13. Suggested design of ramp vehicle to provide electrical power to weapons

1—Turbine prime mover
2—Gear box and driven units such as alternator, generator, and pumps
3—Control space
4—Tanks for various fluids and ballast
5—Air-conditioning unit

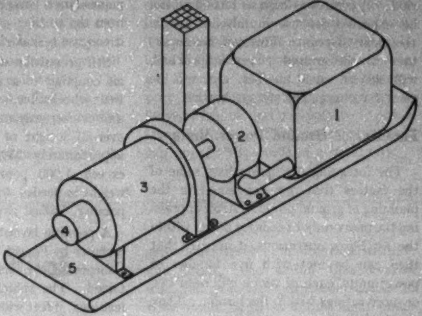


Fig. 14. Suggested arrangement of stationary electric generating set

1—Turbine drive
2—Free piston compressor
3—Alternator and gear reduction
4—Exciter
5—Mounting base

Wave shape of the output voltage will be more accurately limited in the future. The trend in the reduction of any deviation from a sine wave is shown by Fig. 7. These reductions apply to harmonics, single or multiple, and to crest factor. They are of particular importance to the ground power unit since it must match the quality of the air-borne units. Modulation is an undetermined factor in this trend at this writing. Modulation can occur to both voltage and frequency. The variable speed drive is a major source of modulation, particularly in voltage and to a lesser extent in frequency. Oscillations within the drive can vary frequency at about 1 cps per second, and apparently this is not detectable by the instruments that are normally used. Better instrumentation will permit its detection and may also lead to a campaign to reduce the modulation. The matter of exactly matching the quality of air-borne alternators could force the adoption of variable speed drives on ground power units.

The modulation of frequency can cause a considerable pulse of voltage. Measurements by one weapon manufacturer with the co-operation of an electrical supplier have shown these pulses to be more than 10% of line voltage, far exceeding the permitted variation of regulated voltage. Again, these pulses of voltage have not been the designated cause of any difficulty in load equipment, but it is feared that if instrumentation is made available that will detect the pulses, there will follow an effort to reduce or remove the pulses

Recovery from surges caused by the addition or subtraction of unit load is also scheduled to tighten as shown by Fig. 8. A single surge of power in an airborne system is not particularly disturbing but periodic surges can be very disturbing. A single surge will throw the pip of a radar scope off limits, but it soon recovers. Repetitive surges can be troublesome. In missile flight, the electrical loading can be scheduled so that surges are not troublesome.

As more experience is gained with 3-phase electrical systems, more of the individual loads go from single to 3-phase operation. The weapon manufacturers are very careful to balance single-phase loads by distribution among the three phases. With this progress comes a demand for less and less difference in phase voltages. The trend is shown by Fig. 9.

Other Secondary Power

The comparable trends in hydraulic systems will be of interest to the electrical engineer, particularly one who is concerned with ground power. The trend in system pressure is shown by Fig. 10. The greater power demands that are coming in the next 10 years would justify a still higher growth in pressure, but the zone of diminishing returns is reached at about 5,000 psi and even less in smaller aircraft. Beyond this critical value, the weight of heavier tubing offsets potential gains.

The trend in fluid temperature within the hydraulic system is shown by Fig. 11. The growth in temperature is due to two factors; the demand for more power, and the proximity to sources of great heat. Silicone fluids will permit a rise to about 800 degrees Fahrenheit from the current values of 275 degrees Fahrenheit. To achieve the maximum limits that are forecast, it is likely that lessons will be drawn from nuclear experience to use liquid metals as the fluid. Hydraulic fluids face another hazard of the nuclear era. Neutron radiation will damage many of the hydraulic fluids that are known today.

There will be an increasing demand for more accurate control of system pressure, and the possible trend is shown by Fig. 12. This requirement is not as severe as that which has been imposed on the electric system. At the same time that the regulation is reduced, it is likely that suppression or reduction of ripple will be required. As pump operating speed grows from the current rating of approximately 4,000 toward 12,000 rpm in the next 10 years, it is likely that the amount of ripple will decrease.

Pneumatic power will follow the same trends as hydraulic with regard to pressure and ripple in high-pressure air-borne systems. The pressure in stored systems is likely to more than double those for operating systems in the same time period. Pressures as high as 12,000 psi can be expected for use in missiles during the next 10 years. This can be important to the ground power units which will service such missiles, and will be used to recharge the storage bottles.

Families of Ground Power Units

The foregoing discussion lists some of the factors that are concerned in the planning of ground power. After appraising the many major needs for power within the Air Force operations, it appears that they can be met with five families of power units, each of which will need two or more ratings within the family. They are the following:

1. A self-propelled ramp vehicle to supply electric (and other) power for servicing and maintaining aircraft or missiles. This vehicle can take two forms; the first a conventional vehicle that can be transported in cargo aircraft, and the second, a "pod" mounted on but separable from the tactical aircraft that it supports.
2. A trailer-mounted unit for nonperiodic use. These will be single purpose units as for hydraulic power or air conditioning.
3. Portable (skid mounted, stationary) electric generating sets of higher power rating than that of the mobile vehicle.
4. Electric motor drives.
5. Fixed power plants, electric, of rather large ratings to service detection and missile control centers.

The first family has been of greatest interest to the Air Force and to weapons manufacturers. The concept of a ramp vehicle has grown steadily in recent years, and several designs have appeared. Today we are faced with the need for still more power from such a power unit, and we have a chance to do some planning around a turbine prime mover within that vehicle. The turbine is ideally suited to driving an alternator at constant speed, as it is essentially a constant-speed device itself. Since a turbine is used, it gives a capability of delivering low-pressure pneumatic power in considerable amounts, and this bleed air can be used for other purposes. Today this air can be used for pneumatic starters on the weapon engines, for ground testing of free-air or ram-air driven power units within the weapon, and as a source of pressurization and of ventilation, since the output air is harmless.

When electric power is delivered and used within a weapon, the excess heat, virtually all of that power, must be withdrawn by means of cooling air. This means that the power unit must have an air conditioner to generate that cooling air. The air conditioner can use air cycle techniques since compressed air is available from the turbine compressor section, or it can use a shaft-driven compressor.

Self-propulsion will be provided by an air coupling to a separate turbine unit. Four-wheel drive is desirable, because the unit can serve as a tow vehicle also. The over-all weight of a combined unit of approximately 250-hp rating need not exceed 4,000 pounds. If higher dead weight is needed to improve towing capability, ballast can be added.

A possible layout of a suitable vehicle is shown by Fig. 13. It will be noted that a surprisingly low silhouette can be maintained. The vehicle as sketched is 15 feet long, 6 feet wide, and little more than 4 feet high. The secondary power section of the vehicle has been located at the rear of the unit, to make it easier to meet the needs of the weapon system concept. This concept as practiced by the Air Force and the aviation industry has accelerated the development of advanced weapons. It is very jealously guarded by both the military and the industry, and yet both agree that a reasonable amount of standardization is to be desired. For that reason, those elements of the ground power unit that are concerned with matching the secondary power system of a weapon exactly are kept to the rear of the unit. The four wheels, propulsive means, air conditioner, and turbine drive are not concerned with any particular weapon, and are portions of the vehicle that can be standardized. The secondary power alternator (and generator, if used) are directly involved in the matching of airborne characteristics, and they are located to the rear of the vehicle where they can be readily interchanged. Some designs of alternators do not use Army-Navy standard mounting flanges, and it is proposed that the reduction gear box which receives power from the turbine and drives the alternator, also be readily removable. Now, by changing a module at the rear of the vehicle to another similar unit that matches the airborne equipment, an exact matching of the air-borne system can be made. Associated equipment and controls should also be removable through the modular concept of design.

In the sketch of Fig. 13, the space 1 has been allocated for the turbine prime mover. Inlet and exhaust are on the upper surface of the vehicle. Space 2 is for the reduction gear box and driven devices, alternators, generators, pumps, and air-conditioning compressor, as needed. Space 3 is for the control of the secondary power system. Space 4 represents ballast tanks which can carry a variety of fluids. Fuel for the turbine is an obvious need. If ballast is needed, other tanks can be filled with water. A tank can be reserved for hydraulic fluid reservoir if the weapon needs it. If the weapon has a ram-jet engine, one or two tanks can be used as a storage for the simulated fuel that is needed for ground tests of the engine. The air-conditioner unit is located at space 5. By treating the gear box and its generators of power as a removable module, it is possible to design the vehicle to use or to deliver more than one type of secondary power. There is no reason why a hydraulic or pneumatic pump unit cannot be driven at the same time as the electric generator and air conditioner.

In the second family of power units, the first type that is likely to reach power values which justify a turbine drive is the hydraulic power unit. Within 10 years, a unit to deliver 50 gallons per minute at 5,000 psi will be needed. This will require a 200-hp prime mover. It should be designed with the module concept also, because it is apparent that in the rapid development of hydraulic systems, many pumps and fluids will emerge. One weapon may have systems of differing pressure ratings, and systems that use different fluids. After 10 years, there may be a need for external cooling during the ground checking of hydraulic systems, and this will require an air conditioner unit as well.

A suggested form for the stationary power unit of the third family is shown by Fig. 14. Again, the turbine is selected as the prime mover. This type of unit may be the successor to the familiar diesel-engine-driven alternator. As shown, the unit is driving a 400-cps alternator, and this is to be recommended for all Air Force power around weapons and their associated controls. In this type of unit, fuel economy is to be sought, since it will be called on to operate continuously. It is suggested that the unit be able to operate without shutdown for 9,000 hours, or one year, to meet Air Force needs in its remote stations.

A free piston compressor appears to be a promising source of the hot gases to drive this turbine. It can use a great variety of fuels, and with diesel oil, it is able to exceed the performance of the diesel engine in regard to fuel economy. A very conservative design of a 500-kw unit to deliver power at 400 cps should not weigh more than 3,000 pounds. The combination should offer fuel economy, long life between shutdowns and overhaul, and weight low enough to make transportation by air very easy. Among the disadvantages, this compressor is still a reciprocating device, although it is simpler than its diesel engine counterpart. It also needs cooling, usually liquid, which adds to its complexity. Filtering of the inlet air will be just as important as for the reciprocating engine it supplants.

In line with the thought that the stationary power unit deliver 400-cps power, it is also suggested that the last family to be considered in this paper, the electric-motor drives, also be designed to use 400-cps power. They can use higher speeds and this, with the higher frequency, will give great reductions in unit weight. The higher speeds are more suitable to the driving of many of the pump units that are forthcoming.

The fixed power plants for missile detection and control centers, are too large a subject to be a part of this paper. They too will present some notable problems in quality of power, reliability, life, and security. Ground power for the support of weapon systems has grown in 20 years from a unit comparable to power lawn mower to a fixed power plant that can supply the needs of a small city.

Informative Feedback in Jet-Pilot Control-Stick Motion

N. D. DIAMANTIDES
ASSOCIATE MEMBER AIEE

Synopsis: In many closed-loop systems in which a human operator forms a link, the operator adds a special component to the error signal in continuously monitoring the system transfer function. Such a signal component, known in cybernetics as informative feedback, was found to be characteristically present in the control-stick motion of jet-airplane pilots. It was found, furthermore, that the amplitude of this feedback signal is subject to a learning process and that the over-all phenomenon, which falls within the scope of both control engineering and psychology, is amenable to quantitative analysis.

THE DESIGN of any block in a closed-loop servo system is dependent on all the blocks of the system. Wherever a human being is involved, as in a manned control system, the over-all design is generally prejudiced by lack of precise information on the human block. It is in this light that the increasing studies of the human organism in terms of a transfer function or, since nonlinearities are involved, a "describing function," which recent years have witnessed, are fully justified. The possible usefulness of such a concept increases yearly with the difficulty of the human operator's role in handling machines of greater complexity and speed. Rising costs and added safety hazards of such machines also heighten the need for more accurate performance prediction prior to construction.

In the specific case of piloted high-speed aircraft, the availability of an adequate analog of the pilot would provide a highly desirable tool for both airframe and flight-trainer design.

A number of pilot simulators have been built and tried both in the United States and abroad. However, the novelty of the art and limited knowledge of the human organism, viewed as a highly nonlinear and extremely complex mechanism, have hindered the success of simulating the physiological and psychological functions of the prototype. Nevertheless, in combination with accepted statistical methods they constitute probably the only effective means for evaluating the adequacy of the mathematical models constructed to represent analogs of human functions. This paper deals with one specific aspect of the human-analog study: the phenomenon of informative feedback in pilot control behavior.

An airplane simulator designed at Goodyear Aircraft Corporation for pilot control-efficiency evaluations made possible the observation of a highly interesting feedback phenomenon in pilot control-stick hand action. This action was evoked as follows. The pilot was placed in a mock-up cockpit whose horizontal position was disturbed by a simulated wind. The pilot's task was to keep the ship flying level at a constant altitude and speed. The action thus evoked was, for practical purposes, similar to what would occur in actual flight under similar atmospheric conditions. Methodologically the investigation was based on a parallel study of the performance of the human operator and his analog, which was built of electronic components on an analog computer.[1,2] To avoid the possibility of physiological crosstalk, only one degree of motion, pitch, was simulated.

Principles

A neuromuscular action such as the motion of a pilot's hand on the control column must be initiated by a command from the central nervous system. In the pilot and simulator situation such a command was evoked by the amplitude and rate of motion of the artificial horizon and by the acceleration of the seat. These visual and motional stimuli informed the pilot that a disturbance was taking place that he was to correct by making the proper control-stick motion.

Such a signal flow from the sensor units through the central nervous system to the motor units is influenced by the two well-known physiological discontinuities of threshold and reaction time. These two inherent characteristics of the human servo act as a gate in the path of the information transfer between input and output.

The input signal received by this servo is further modified by other characteristic parameters of the system. Such parameters are attributable to both the neurophysiology of the central nervous system, and the nature of the sensor and motor elements. Their effect is a proportional and derivative one (rate). For instance, when the apparent movement of the actual or artificial horizon indicates airplane displacement, the eyes send a signal to the brain that is related to the amplitude of the displacement. Acceleration will cause a signal to be generated in the vestibular organs of the

---

Paper 57-494, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, March 7-9, 1957. Manuscript submitted July 24, 1956; made available for printing February 21, 1957.

N. D. Diamantides is with the Goodyear Aircraft Corporation, Akron, Ohio.

The airplane mock-up and the original version of the pilot simulator were designed by R. J. Mead, who is now with McDonnell Aircraft Corporation. The mock-up was contractually supported by the U. S. Navy, Bureau of Aeronautics. An improved version of the simulator, a joint effort of the author and R. J. Mead, was used in this study. Dr. A. J. Cacioppo introduced the psychological approach to the problem. It should be noted, however, that the writer accepts sole responsibility for the conclusions reached in this paper.

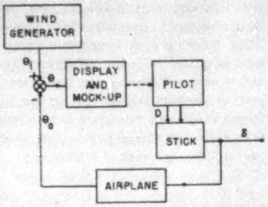


Fig. 1. Block diagram of closed-loop system

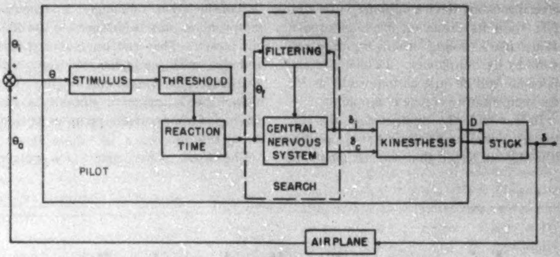


Fig. 2. Block diagram of physiological closed-loop system

inner ear and the proprioceptors of the muscles, tendons, and joints. Rate information is provided by the ear canals through endolymph flow, and by the eyes through peripheral vision resulting from the successive stimulation of the rods of the macula. In the case of the motor elements, it is known that the muscles involved contract or relax and at the same time move with their supporting bones in carrying out the commands of the brain. Consequently, it is seen that displacement and its time derivatives are important as causes of inertial forces, muscle elasticity, and fluid damping.

Apparatus

The apparatus used in the study of the feedback occurrence in question was simply a pilot seat mounted rigidly on a platform hydraulically driven about a horizontal axis. A visual display and a control column were placed on the platform. The display consisted of a 5-inch laboratory cathode-ray tube masked to the size and shape of a gyrohorizon instrument. The trace was deflected vertically as the platform was moved. In this manner the movement of the seat during actual airplane pitch as well as the related movement of the artificial horizon could be represented. Only pitch-control investigation was attempted to restrict the measurements to a single degree of freedom.

A 24-inch length of pipe having an actual airplane-handgrip served as the control column. The column could be moved about a horizontal axis located under the platform floor and was made to work against an inertial load, an adjustable spring, and viscous-friction forces. A damper cylinder with a control valve produced the viscous damping. Control-stick motion was translated into a proportional electric signal by a potentiometer, which was located under the platform. This signal was fed back into the mock-up and display to overcome the original disturbance.

Fig. 1 shows the closed-loop system as formed by its four basic components of mock-up and display, pilot, stick, and airplane. Since the three nonhuman components were defined in advance, the problem was to study the make-up of the pilot block. The role of input *D* is discussed in the appendix.

For the study, an electronic analog of this physical loop was constructed of operational and recording units of a GEDA* analog computer. In the analog the display's function of supplying the stimuli was performed by a simulator. Electronic analogs of the control column, the airplane, and the human operator, the latter perceiving and correcting for disturbances, were incorporated. The fact that the task of this study was simple enough as to be practically automatic made this last device feasible.

Fig. 2 translates Fig. 1 into physiological terms. The basic principles involved are discussed in the appendix. It should be noted that random noise effects are neglected and that the pilot block includes search in addition to the basic processes of stimulus, threshold, reaction time, and kinesthesis (neuromuscular action).

Search was introduced to enable the simulator to match the performance of jet-interceptor pilots. When a pilot was called upon to keep his plane flying level against random pitch disturbances caused by turbulent air, his stick motion contained a very pronounced oscillatory component, around 1.4 cycles per second, which at first was present even in periods of quiet atmosphere. The oscillation, or dither, phase-modulated to some extent by the main stick motion, persisted as long as the pilot anticipated pitch disturbances. When this dither was brought to the attention of the pilots at the end of the test series, they attributed it to the fact that they wanted to be certain of having continuous control of the aerodynamic situation, so as to be able to handle the next disturbance. According to the pilots, this type of control is particularly accentuated in situations where they are constantly under tension, e.g., formation flying and landing approaches.

Fig. 3 shows a partial recording of an experienced jet-interceptor pilot's test run. Channels 1 and 2 of the figure record air turbulence and the resultant airplane pitch angle, respectively. Channel 3 records the pilot's corrective motions. The oscillatory motion is obvious in this last channel, superimposed on what, without it, would have been a motion closely resembling that of channel 2.

The frequency of the oscillatory motion is of considerable interest in itself. In an ideal case, where the pilot attempts to follow a step input while the control loop is open, i.e., $\theta_o = 0$ (see Fig. 2), and the step amplitude is much greater than the threshold level so that the threshold effect need not be considered, the two parameters characterizing the response will be the previously mentioned reaction time, τ, and the muscular-lag time constant, τ_0. The latter being the predominant kinesthetic factor.

When typical values for the two time constants, $\tau = 0.25$ second and $\tau_0 = 0.125$ second,[3] are used and the pure gain factor is disregarded, a transfer function between input (stimulus) and response (control-force output of the arm) can be presented as the following equation

$$G(s) = \frac{e^{-0.25s}}{0.125s + 1} \qquad (1)$$

where s is the Laplace operator.

---

* GEDA is a registered trade-mark of Goodyear Corporation.

Vo
Pt.
19

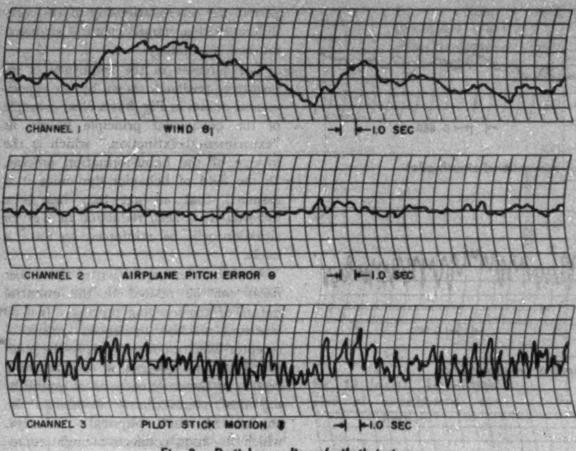


Fig. 3. Partial recording of pilot's test run

In the frequency domain, the reaction time factor, exp (-0.25ω), becomes

$$e^{-j0.25\omega} = \cos(0.25\omega) - j\sin(0.25\omega) = 1\underline{/\phi = -0.25\omega} \quad (2)$$

which means that the phase shift, ϕ, is proportional to the frequency. Similarly, the muscular-lag factor, $(0.125s+1)^{-1}$, becomes

$$(1+j0.125\omega)^{-1} = \frac{1-j0.125\omega}{1+(0.125\omega)^2} = \frac{1}{\sqrt{1+(0.125\omega)^2}}\underline{/\phi_0 = -\tan^{-1}(0.125\omega)} \quad (3)$$

Therefore, the total phase shift is

$$\Phi = \phi + \phi_0 = -[0.25\omega + \tan^{-1}(0.125\omega)]$$

in radians. When $\omega = 2\pi f$ and $\Phi^\circ = 180/\pi \;\; \Phi$ are substituted, the function can be presented as

$$\Phi^\circ = 57.29[1.570f + \tan^{-1}(0.785f)] \quad (4)$$

The plot of Φ° versus frequency in cycles per second of Fig. 4 shows that the phase reversal, $\Phi^\circ = 180$, occurs at about $f = 1.5$ cycles per second which is practically the center of the observed dither frequency range.

Informative Feedback

In the study conducted, the presence of dither was unmistakably observed during the performance of jet pilots. Its cause can be looked for in the way high-speed aircraft have to be flown. Piloting depends on the knowledge of the ship's performance as influenced by its high speed, that is, on the knowledge of the performance characteristics of the plane-air system. In continuously acquiring this knowledge the pilot imparts a succession of small fast impulses to the control column. Though these do not deflect the airplane from its course, they are quite sufficient to give the pilot a kinesthetic sense of whether the plane is in danger of getting out of hand. This source of pilot awareness and its control effects are known in cybernetics as informative feedback.[4]

The relationship between feedback and control output permits a planned variation of what in servo terminology would be called the pilot's transfer function. The pilot superimposes a weak high-frequency input, δ_i, on the control signal δ_c, and takes from input θ a partial output, θ_f, of the same high frequency separated from the rest of θ through an appropriate filtering process; see Fig. 2. He explores the amplitude-phase relations of the high-frequency output component, θ_f, to the input δ_i, to obtain the performance characteristics of the plane-air system. What he learns allows him to modify his own transfer function for optimum control, very much in the manner of a self-adjusting servo.

The advantage of this type of feedback is that the resulting adjustments can give stability for every type of constant load within the limitations of both operator and machine. If the characteristics of the load change slowly enough in comparison with the changes of the original input and if the reading of the load condition is accurate, the system is prevented from going into instability. It is assumed, of course, that the load characteristics at high frequencies are the same as, or at least give a good indication of, the load characteristics at low frequencies.

The aspect of the plane-air system that chiefly interests the pilot of a high-speed airplane is the g force caused by changes in the velocity vector that occur when the stick is deflected or the throttle setting changed. A high g force endangers both the airplane and pilot. The dither observed in the performance of the jet-interceptor pilots who cooperated in this human-dynamics study at Goodyear Aircraft Corporation was apparently the informative feedback necessary to inform the pilot as to how many g's he could safely pull at any moment and, therefore, how great a stick deflection he could afford without danger.

From the flight-simulation results pilots could be placed into three categories: those who applied no dither, those who did, and those who did but learned not to after repeated runs. The subject for whom no informative feedback was apparent (see Fig. 5) had nothing but conventional aircraft experience. All the jet pilots carried the informative-feedback habit of their normal flight experience to mock-up flight, see Fig. 6. One of the jet pilots exhibited no tendency

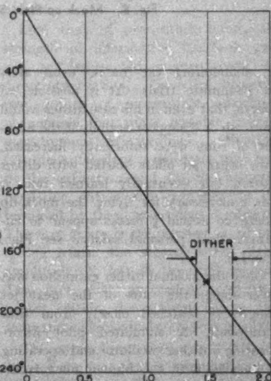


Fig. 4. Phase shift versus frequency in open-loop system

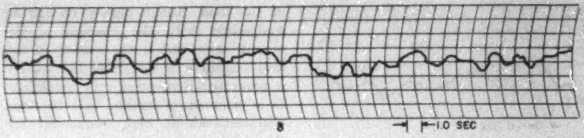


Fig. 5. Mock-up flight test record of conventional-aircraft pilot

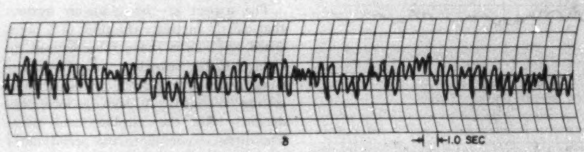


Fig. 6. Mock-up flight test record of jet-aircraft pilot

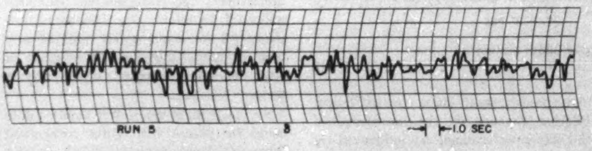


Fig. 7. Mock-up flight 5 showing presence of dither habit

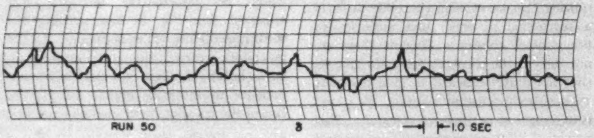


Fig. 8. Mock-up flight 50 showing loss of dither habit

of abandoning the habit, even after 80 10-minute trials. It is possible, of course, that even in his case dither would diminish or disappear entirely if the number of runs were sufficiently increased. Two other jet pilots started with dither motion but eventually learned that it was unnecessary for flying the mock-up where no actual g forces appear to endanger their personal safety, see Figs. 7 and 8.

Once the gradual dither extinction was ascertained, the runs of the next jet pilot were watched closely from that standpoint. A simulated pilot incorporating a dither oscillator and operating in parallel with the human pilot facilitated day-to-day measurements of the dither present in the subject's stick motion. Fig. 9 shows the gradual elimination of informative feedback as the subject's subconscious awareness of its uselessness increased. The experimental data represented in the figure could become the nucleus of a study in itself. The points on the plot show the dither amplitude for the last run of each successive day. It should be emphasized that these points were determined by visual comparison of the stick motions of human and dummy pilots operating in parallel and for the same θ input. No statistical approach was attempted.

Experimental Extinction of Dither

The gradual elimination of dither is a study to which psychological learning theory may be quantitatively applied. Since it is also a control-process phenomenon it is worth further analysis if for no other reason than that it may permit the interrelationship of two very different fields, namely, control engineering and psychology.

The decrease in the amplitude of informative feedback shown by the experimental points in Fig. 9 is a clear example of the behavioral principle known as "experimental extinction," which is the inverse of the habit-formation process. In the field of behavior dynamics it is assumed that an act that has been followed by a need reduction in a given situation will be reinforced to form a habit.

Actually, the acquisition of the dither habit can be related to the essential nature of the learning process. Just as inherited reaction tendencies consist of receptor-effector connections, so the process of learning consists in strengthening certain of these connections or establishing new connections.[5] It should be emphasized that physical pathways, which the term connection might represent to an electrical engineer, are not here implied. What is meant, ultimately, is a functional association between cause and effect. Neither psychology nor physiology has yet provided conclusive proofs as to what the physical representations of these associations are or where they are located in the brain.[6] Furthermore, while organic evolution rigorously tends to eliminate redundant organs or their functions, it often makes a single organ adaptable to a number of functions.[7] Thus, the nervous system is sufficiently plastic to allow (although not always) the transfer of learned reactions. For these reasons it becomes questionable whether such reactions can be sufficiently localized at present to permit accurate physical representation.

In controlling a high-speed aircraft, inherited tendencies cannot be considered dependable, since the human machine was never meant to fly by its own power nor to participate in sonic or supersonic motions with its own unaided controls. Therefore, a great many of the adaptive acts necessary for the undertaking of this biologically strange adventure must be learned. The use of informative feedback is quite patently such a learned act. It is the outcome of the process of habit formation.

In a jet-interceptor pilot the habit is probably acquired as follows. Confronted with the task of controlling the vehicle either for the first time or under tense conditions, the pilot does not possess as part of his evolutionary endowment the means of direct and instant "feel" of the aerodynamics involved. Yet he knows that these aerodynamics

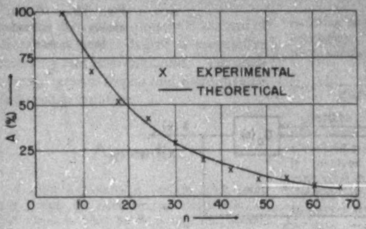


Fig. 9 (left). Elimination of dither by experimental extinction

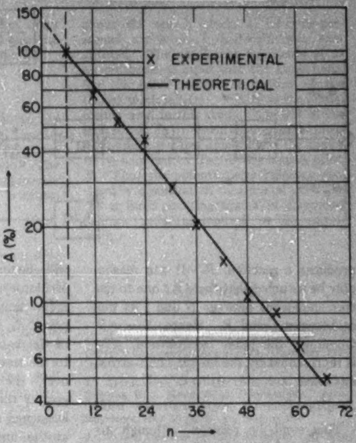


Fig. 11 (right). Logarithmic plot of experimental extinction

are varied continuously by the interaction of such factors as airplane characteristics, air speed, and air density and that his survival depends on his ability to alert himself to every change.

Apprehension of physical injury is the drive that evokes pilot action. Under the circumstances, dither is the action best suited to gain the knowledge required to allay his apprehension by providing continuous information on the plane-air system. Although the innate stimulus originating dither might be entirely casual at first, the dither reaction takes place in temporal contiguity with a stimulus energy that is followed by a diminution in the drive. That is, the output motion filtered and compared with the initial dither input provides the reassuring information being sought. The hypothesis of primary reinforcement suggests, therefore, that the tendency of the original stimulus to occur is likely to be incremented on subsequent occasions. Thus, a beneficial habit is gradually established and the threat to survival diminished accordingly.

In mock-up flying the informative feedback stimulus is not actually critical. The pilot, however, has no inner monitor to tell him in advance which stimuli, simple or complex, are associated with the critical causal factors of his dither reaction. Inevitably he reacts to the reinforced stimulus element associated with the dither.

Although the dither learning process is characterized by the setting-up and strengthening of specific receptor-effector connections, its effect is not limited to reinforcing these connections only. Connections are in fact mediated between a great many other receptor-effector processes. By no means are all the factors of any reaction situation critical to the extent that their presence is required in bringing about need reduction. In other words, the reaction involved in the original conditioning becomes connected with a considerable zone of stimuli other than, but adjacent to, the stimulus conventionally involved. This phenomenon, known as stimulus generalization, was certain to manifest itself in the experiment under discussion where the strong environmental similarities between actual and mock-up flying were bound to result in the overlapping of stimulus continua.

The mock-up contrasts with the plane-air system in one important respect: its transfer function remains fixed. Consequently, stimuli that previously evoked the informative feedback are divested of causal relationship to the critical factor in the reinforcement process. Although these stimuli continue to evoke the dither reaction, the reaction is not followed by reinforcement. Since the dither effort becomes wasteful and unadaptive, its gradual abandonment can be expected. The psychological axiom upon which this hypothesis is based is that a habit not followed by reinforcement can invariably be forced into experimental extinction.

A habit, functionally considered, is formed by receptor-effector connections set up by reinforcement. Likewise, the process of habit extinction consists of physiological subtraction of a series of discrete decrements, each decrement resulting from a distinct receptor-effector conjunction closely associated with the process.

Like that of any other habit, the strength of informative feedback cannot be determined quantitatively by direct observation. There are, however, two groups of measurable associated phenomena that make possible indirect measurement of feedback extinction: (1) the antecedent conditions that lead to the extinction and (2) the behavior persisting within the organism as determined by these antecedent conditions. The first group is the number, n, of preceding trials; the second is the dither amplitude, R, left in the stick motion. The fact that antecedent conditions persist within the system means that some sort of biological storage is involved.

The symbolism, $\dot{S} \to \dot{s} - - \to r \to R$, expresses the fact that a stimulus energy, $\dot{S}$, acting on a receptor, initiates an afferent receptor discharge, $\dot{s}$, that initiates an efferent discharge, r, which in turn enters the effector system and

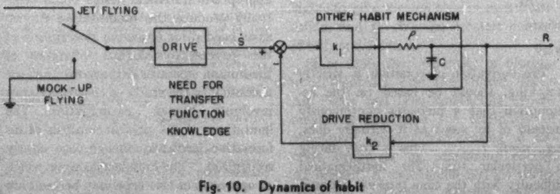


Fig. 10. Dynamics of habit

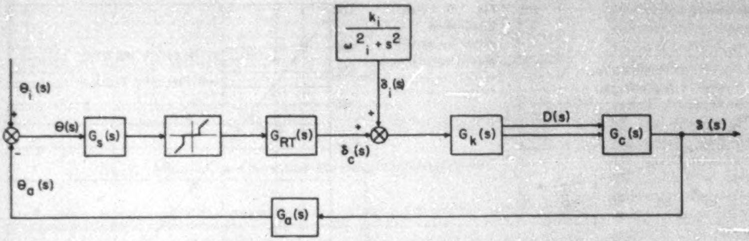


Fig. 12. Actua analog of closed-loop system

produces a reaction, R. It can reasonably be assumed that there is a one-to-one correspondence between $\dot{S}$ and $\dot{s}$ as well as between r and R. Hence, $\dot{S} \dashrightarrow R$. Obviously the habit organization, ${}_{\dot{s}}H_R$, is represented by the broken-shaft arrow, and thus

$$\dot{S} \rightarrow {}_{\dot{s}}H_R \rightarrow R \tag{5}$$

This symbolic relation, although accepted in psychology, does not convey the notion of habit reinforcement. The foregoing discussion makes it plain that, from the standpoint of control engineering, habit formation and extinction are a closed-loop process. The habit mechanism is represented by the forward branch of the loop and the drive reduction by the feedback. Besides the aforementioned storage element, c, which is present in the mechanism, a dissipative element, ρ, should be included in view of the rather low efficiency characteristic exhibited by all biophysical functions. Consequently, in the way of a very rough approximation, the dynamics of the habit can be modeled as shown in Fig. 10, where the gains, k_1 and k_2, measure, respectively, the vigor with which the habit is acquired and the drive reduced. The response of the loop when the switch is in the JET FLYING position represents the formation of the dither habit. With jet training considered a step input to the drive complex of the pilot and with the transfer function of the foregoing oversimplified habit dynamics known from the diagram to be

$$R = \frac{k_1}{1+k_1k_2}\,\frac{1}{\dfrac{\rho c}{1+k_1k_2}s+1}\left(\dot{S}+R_0\,\frac{1+k_1k_2}{k_1}\right) \tag{6}$$

several facts can readily be observed. R_0 is the value of R at the beginning of the experiment. For instance, in the habit-formation stage the habit strength increases exponentially up to an upper physiological limit, beyond which no further increase is possible. The nature of the informative feedback suggests that there is a maximum value of dither amplitude, beyond which the airplane will begin to show an objectionable oscillatory motion. Clearly, this maximum amplitude will be the starting point, R_0, for the experimental extinction of the habit when the switch is thrown to the MOCK-UP FLYING position. Moreover, the greater k_1 and k_2 are, the greater the speed with which the dither is learned and abandoned. Factors such as the unconditioned stimulus and the amount of work involved in the habit mechanism influence the magnitude of the gains. It should be noticed that when the inverse Laplace transform of equation 6 is taken, the dither amplitude, R, is expressed as a function of the time, t, that has elapsed between the start of the experiment and the present. From a psychological standpoint, though, this time is effectively equivalent to the number, n, of trials performed during t. This is especially so when the trials are similar and of equal duration, τ_x. Hence, in the present experiment, $t \rightarrow \beta n \tau_x$. The parameter β depends on the design of the experiment. The greater the intertrial interval is, the smaller β is. Thus, the physical solution of equation 6 for mock-up flying, $\dot{S}=0$, becomes

$$R = R_0\epsilon^{-an} \tag{7}$$

where

$$a = \frac{(1+k_1k_2)\beta\tau_x}{\rho c}$$

Since the data of Fig. 9 were plotted in per cent of the sixth run, $n_0 = 6$, $R_0 = R_6$, we find that the normalized dither amplitude, A as a function of the number of trials will be

$$A = \frac{R_n}{R_6}\,\epsilon^{-a(n-6)} \tag{8}$$

where

$$n \geqq 6$$

The derivation of equation 8, strictly speaking, should be based on the assumption that n increases continuously instead of by steps as it actually does, e.g., $n = 6, 7, \ldots$. This is true during a particular run. The mathematical difficulty with the time interval between runs can be overcome if it is postulated that parameter β compensates for the effect the interval might otherwise have on the learning process. To calculate the constant, a, the experimental data were plotted on logarithmic paper (see Fig. 11) and, from the slope of the straight line best defined by the data, the value of a was derived as $a = 0.0496$. Thus,

$$A = \epsilon^{-0.0496(n-6)} \tag{9}$$

A plot of equation 9 is given by the solid line of Fig. 9.

From equation 7 we find

$$(\Delta R) = -aR(\Delta n) \tag{10}$$

That is, the decrease in dither amplitude is proportional to both the number of antecedent trials, Δn, and the amplitude, R, at the beginning of Δn. This is a rather common empirical fact in many extinction experiments. Also significant is the fact that the model of habit dynamics, shown in Fig. 10, provides for the investigation of the relations between the response and each independently manipulatable environmental variable as well as the manner in which these variables interact. From an over-all point of view, equation 9 relates an actual instance of habit extinction to the theory that experimental psychologists have adduced on the basis of laboratory experiment.

Discussion

The results outlined indicate that analog simulation of psychological phenomena in real time could be a powerful tool in conducting controlled tests on the validity of models based on the biophysics of the human organism. Such a tool with an appropriate statistical design could eventually produce the quantitative expressions needed for objective description of the psycho-physiological activities of the human operator-expressions that are necessary to the study of control systems involving human participation. The limitations of the present analysis of informative feedback, though, are clearly recognized. The foundations upon which an approach of this type may be built are none too firm at the present time, and of necessity a certain degree of speculative reasoning has been employed. The analysis will require ampler and more diversified experimentation.

Appendix

A pilot associates a certain stick force per g with the operating characteristics of the airplane that he is flying, i.e., speed, altitude, weight, and center of gravity. Therefore, the flight path attitude, γ, could be selected as the primary forcing variable, since the stick-load-factor increment is $V\dot{\gamma}/g$. In the simulated situation, however, the mock-up stick force was not affected by the flight path attitude since aerodynamic forces on control surfaces were absent. Consequently, the pitch attitude, θ, had to be selected as the primary variable in supplying the pilot with response cues. In other words, a function of θ constitutes the stimulus complex initiating the response.

In general terms, the transfer function for the stimulus block, $G_s(s)$, was based on the fact that the pilot is perceiving signals proportional to the actual airplane pitch, $a_0\theta$, and its first two derivatives, rate, $a_1s\theta$, and acceleration, $a_2s^2\theta$. Consequently,

$$G_s(s)=a_2s^2+a_1s+a_0$$

The control-column transfer function, $G_c(s)$, is the familiar second-order one of a system comprising inertia, I, compliance, $1/k$, and viscous damping, c:

$$G_c(s)=Is^2+cs+k$$

The airplane transfer function, $G_a(s)$, for the pitch-angle response, θ_a, to elevator motion, δ, is fairly simple since only one input is involved. For rapid tracking motions it can be further simplified by omitting the phugoid, or long-term pitch oscillation. Thus, for the particular jet plane simulated, this transfer function is

$$G_a(s)=\frac{K(\tau_1s+1)}{s(\tau_2^2s^2+2\zeta\tau_2s+1)}$$

For the building blocks involved in the pilot simulator, no Laplace form for the threshold transfer function is possible since it involves a nonlinear operation. The pure time delay, however, which represents the physiological reaction time, τ, is known to have a transfer function of

$$G_{NT}(s)=\epsilon^{-\tau s}$$

The kinesthetic operation representative of the biomotor mechanism of arm and hand again involves elasticity, damping, and inertia in limb, muscle, and ligament. Its transfer function, therefore, can be given to a first approximation by a second-degree polynomial, $k'/(ms^2+ns+1)$, if open-loop performance is assumed. However, in a normal organism such performance will lead to ataxia, a cessation of communication between effector elements and the central nervous system. In this state the effectors, while remaining healthy, are no longer monitored. But such open-loop performance is not the case in a normal person. As a matter of fact, the proprioceptors throughout the muscles and joints of the body continuously monitor the effector organs. Their feedback signals are subject to a transmission time lag that can be approximated by an operation $k_0/(1s+1)$ in the feedback branch. Thus, the over-all kinesthetic transfer function becomes

$$G_k(s)=\frac{k'}{1+k_0k'}\times\frac{ls+1}{\frac{lm}{1+k'k_0}s^3+\frac{m+nl}{1+k'k_0}s^2+\frac{n+1}{1+k'k_0}s+1}$$

Finally, since virtually nothing is known about the search block, a 1.4-cycle-per-second oscillator, $k_t/(\omega_t^2+s^2)$, was substituted. The oscillation represents δ_t as an independent input into the kinesthesis. On the other hand, the output of the $G_{NT}(s)$ block is $\delta_e(s)$. Fig. 12 is the actual analog used in this study. Control of the gain factor, k_t, by potentiometer adjustment permitted the dither amplitude of the analog to be matched to that of the human pilot, R. Thus the potentiometer readings provided a quantitative measure of R. The dither force resulting from δ_t is illustrated in Figs. 1 and 2 as input D to the control stick.

References

1. INVESTIGATION OF CONTROL "FEEL" EFFECTS ON THE DYNAMICS OF A PILOTED AIRCRAFT SYSTEM. *Goodyear Engineering Report-6726*, Goodyear Aircraft Corporation, Akron, Ohio, 1955.

2. APPLICATION OF GEDA TO HUMAN DYNAMICS STUDIES, R. J. Mead, N. D. Diamantides. *Aeronautical Electronics Digest*, Institute of Radio Engineers, New York, N. Y., 1955, p. 200.

3. FLIGHT CONTROLS HUMAN DYNAMIC RESPONSE STUDIES. *Report AE-61-6*, Bureau of Aeronautics U. S. Navy, Washington, D. C., Nov. 1953.

4. CYBERNETICS (book), N. Wiener. John Wiley & Sons, Inc., New York, N. Y., 1955, p. 133.

5. PRINCIPLES OF BEHAVIOR (book), C. L. Hull. D. Appleton-Century Company, New York, N. Y., 1943, p. 258.

6. SIGNIFICANCE OF INFORMATION TO NEUROPHYSIOLOGY, J. A. V. Bates. *Transactions*, Professional Group on Information Theory, Institute of Radio Engineers, New York, N. Y., Feb. 1953, p. 140.

7. POSSIBLE FEATURES OF BRAIN FUNCTION AND THEIR IMITATION, W. Grey. *Ibid.*, p. 134.

8. THE KENTUCKY SYMPOSIUM ON LEARNING THEORY, PERSONALITY THEORY AND CLINICAL RESEARCH (book). "Interpretations of Learning Data and Some Recent Developments in Stimulus-Response Theory," K. W. Spence. John Wiley & Sons, Inc., pp. 3–4.

Flat Work-Coil Design

B. E. MATHEWS
NONMEMBER AIEE

IN RECENT YEARS, aerodynamic heating, or air-friction heating, has become one of the major problems in increasing aircraft velocities. Since little is known concerning the effects on aircraft structures heated in this fashion, experimental data must be obtained which will indicate necessary design procedure. It is desirable that the experiments be conducted under laboratory conditions to eliminate the necessity for flight, to give greater ease in instrumentation, and to give the greatest control of test conditions.

Since 1950, the University of Florida, under contract with the United States Air Force, has been conducting experiments to determine the feasibility of simulating aerodynamic heating of aircraft structures by induction heating. At present, a commercial 200-kw radio-frequency induction heating unit is being used as the primary power source. The results of these experiments will be applied to a test facility that will simulate both air friction heating and aerodynamic loading of aircraft structures. This paper is concerned with an analysis of the electrical characteristics of the work coil used to couple power into aircraft structures.

Description of Work Coils

Using induction heating to simulate aerodynamic heating of aircraft structures imposes unique requirements on work coils and their design not usual in the more common uses of radio-frequency heating. Among the more important considerations are:

1. Surface materials are generally nonmagnetic and have low electrical resistivity. Therefore close coupling is necessary to give good efficiency.
2. To simulate effectively the spatial power

Paper 57-493, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, March 7–9, 1957. Manuscript submitted March 5, 1956; made available for printing February 21, 1957.

B. E. MATHEWS is with the University of Florida, Gainesville, Fla.

The author wishes to thank Prof. Roger W. Sampson and Fred R. Sias of the Electrical Engineering staff at the University of Florida for their constant encouragement and for their invaluable assistance in preparing this manuscript.

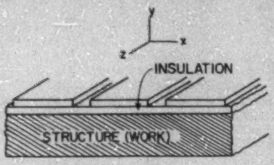


Fig. 1. Cross section of assembly

distributions resulting from aerodynamic heating, surface power should be as uniform as possible.

3. In a combined test using simulation of aerodynamic loading, much of the structure surface may be covered by pressure pads. Coils must fit under these pads.

4. Surface power should not change with warping of the structure.

Experience has indicated that the work coils consisting of wide, flat conductors lying close to the structure give good results. These coils have efficiencies which approach the maximum, produce fairly uniform power distributions, will fit beneath pressure pads, and when under pressure will follow changes in surface contour. In present testing, typical conductors may vary in width from one inch to 1/4 inch depending on the power density required and the current available. In most cases the coils are cut in one piece from copper sheet. Fluid cooling is not generally used and, since sufficient air cooling is difficult to obtain, the thermal capacity of the coil largely determines the amount of power which can be transferred to the structure without melting the coil. The thickness of the conductors is determined by a compromise between ability to follow changes in surface contour and thermal capacity. Generally, conductors with a thickness in the order of 0.050 inch are used.

Because of the voltage drop across the work coil, the coil is separated from the structure by sheets of insulation. Since close coupling is desired, the insulation thickness is kept to a minimum as determined by the voltage between the coil and the structure. A cross section perpendicular to the conductor axes of a complete assembly is represented in Fig. 1.

There are two general types of coils: the zig-zag and the spiral as shown in Fig. 2. The important difference between these is the relative direction of current flow in adjacent conductors and the resulting difference in the magnetic-field configuration. Applying a voltage across the zig-zag coil will cause a current which will be flowing in opposite directions in adjacent conductors. In the spiral coil the current will always be flowing in the same direction in adjacent conductors. These two types of coils have different electrical characteristics due to the difference in field configurations. The zig-zag coil and the spiral coil will, in general, take the form shown in Fig. 2 when heating one side of a structure. If two sides of a specimen are to be heated (or if the specimen is a figure of revolution), the spiral coil may take the form of a helix and encircle the specimen.

Method of Solution

Potential Analog

The work-coil characteristics (such as power density transferred, efficiency, or inductance) are determined by the distribution and magnitude of the fields produced by the work-coil currents. Therefore the work-coil analysis must begin with a description of these fields.

In general, if conductors are carrying current and are surrounded by an infinite, uniform dielectric and if the system is lossless, then the transverse electromagnetic (TEM) wave is propagated. In a TEM wave the electric and magnetic field are everywhere orthogonal and lie in the transverse plane (plane perpendicular to direction of propagation, or in this case, perpendicular to the axes of the conductors). The electric and magnetic fields are solutions of Laplace's equation and must satisfy certain boundary conditions at conductor surfaces.

In induction heating, the fields are not those of a true TEM wave because the conductors and dielectric are not perfect. However, a close approximation to a TEM wave exists when the following two conditions are met:[1]

1. Displacement currents in the conductors are negligible compared to the conduction currents.

2. The intrinsic impedance of the dielectric is much greater than the skin effect surface resistivity of the conductor.

Most metallic conductors and dielectrics in common use satisfy these conditions.

In practice, the region surrounding the conductors consists of a solid dielectric on which the current-carrying conductors are placed, and an air region or other dielectric above the conductors. Since this involves a composite dielectric, a pure TEM wave cannot be supported. Nevertheless, since in the flat coils with close coupling most of the field is concentrated in the dielectric between the work coil and work, the assumption of a single infinite dielectric introduces little error.[2]

The foregoing discussion leads to the conclusion that the field distribution produced by the work coils is, to a close approximation, the same as the distribution of a TEM wave. Also, the electric field produced by static electric charges is a solution of the same equation (Laplace's equation) and satisfies the same boundary conditions as the electric field in the TEM wave. Therefore the same electric field will be produced as in the TEM wave if the conductors are assumed to have static charges rather than to carry currents. In the construction of the electrostatic analogy the relative signs (positive or negative) of the charges correspond to the relative direction of the currents in the conductors, i.e., if adjacent conductors have currents which always flow in the same relative direction (spiral coil) the conductors will be given static charges of the same sign. If adjacent conductors have currents with opposite direction of current flow (zig-zag) the conductors will be given static charges of opposite signs. The electrostatic analogy will be used in this analysis to aid in visualizing the fields

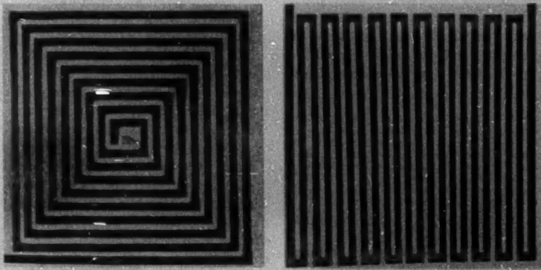


Fig. 2. Two general types of coils, zig-zag and spiral

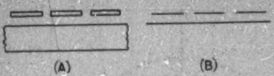


Fig. 3. Assumption of zero-thickness conductors

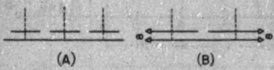


Fig. 4. Assuming wide strips

and because the analytic method used in the solution is more conveniently applied to the static case.

It should be understood that this electrostatic (or potential) analogy is used purely for convenience in describing the spatial distribution of the electric field in the TEM wave. In the actual case the fields are varying with time as the current varies, in this case sinusoidally. At any instant of time, however, this time-varying field has the same spatial distribution as the potential analogy. The potential analogy is used to describe the relative spatial field distribution but the magnitude and effect of this field will be based on the true time-varying conditions. The magnetic-field distribution follows from the electric field distribution since in a TEM wave the two fields are everywhere orthogonal and their magnitudes are related by the intrinsic impedance of the medium.

Now, consider an assembly cross section as that shown in Fig. 1. For the purpose of analysis it will be assumed that there are no field variations in the z-direction. This reduces the problem to a 2-dimensional case in the plane perpendicular to the direction of current flow. This assumes that the fields in the transverse plane are not being influenced by variations of the configuration in the z-direction such as corners where the conductor strips change direction, abrupt changes in the width of the strips, and variations in insulation thickness. The analysis on the basis of the 2-dimensional case will be the solution for all areas which have this cross section and which are not influenced by nearby discontinuities.

SOLUTION PROCEDURE IN ELECTROSTATIC POTENTIAL PROBLEM

The field distribution of the 2-dimensional electrostatic potential analog can be described with the aid of the theory of functions of a complex variable.[3] This method results in general equations that indicate broad design principles.

First a representation of the 2-dimensional configuration is drawn in the z-plane ($z=x+jy$). The conductors are then given appropriate potentials which will result in the desired field distribution as determined by the currents in the actual TEM wave. The next step is to determine the complex-potential function $P(z)$ which describes the field distribution This complex-potential function is an analytic function defined as

$$P(z)=\Phi(x,y)+j\Psi(x,y) \tag{1}$$

where curves of constant Φ are equipotential lines, and curves of constant Ψ are flux lines. Both Φ and Ψ satisfy the Cauchy-Riemann and Laplace equations.

The electric field in the z-plane is given by

$$E(z)=E_x+jE_y=-\left(\frac{\partial\Phi(x,y)}{\partial x}+j\frac{\partial\Phi(x,y)}{\partial y}\right) \tag{2}$$

The total derivative of the complex-potential function is defined as

$$\frac{dP(z)}{dz}=\frac{\partial\Phi(x,y)}{\partial x}-j\frac{\partial\Phi(x,y)}{\partial y} \tag{3}$$

Therefore from equations 2 and 3

$$E(z)=-\overline{\frac{dP(z)}{dz}} \tag{4}$$

where the bar indicates the complex conjugate. Returning to the actual TEM wave, the magnetic field $H(z)$ can be determined from the electric field since the two fields are orthogonal and related by the intrinsic impedance of the medium. The relationship is

$$H(z)=\pm\frac{j}{\eta}E(z)=\pm\frac{j}{\eta}\overline{\frac{dP(z)}{dz}} \tag{5}$$

where η is the intrinsic impedance. The sign of equation 5 depends on an arbitrary choice of the direction of power flow.

Equation 5 indicates that if the complex-potential function $P(z)$ is known, the magnetic- and electric-field distributions can be determined. For many simple configurations, the function $P(z)$ can be written by inspection. However, for complicated cases a transformation procedure can be used. Consider the configuration under study drawn in the z-plane. By the Swartz-Christoffel transformation, the configuration in the z-plane is transformed into the w-plane by a relation

$$z=F(w) \tag{6}$$

where $z=x+jy$, and $w=u+jv$. The resulting configuration is drawn in the w-plane, and the complex-potential function

$$P(w)=\Phi(u,v)+j\Psi(u,v) \tag{7}$$

which describes the fields in this plane can be determined by inspection. Due to the conformal transformation

$$\frac{dP(z)}{dz}=\frac{dP(w)/dw}{dz/dw} \tag{8}$$

Therefore from equations 5 and 8

$$H(z)=\pm\frac{j}{\eta}E(z)=\pm\frac{j}{\eta}\frac{dP(w)/dw}{dz/dw} \tag{9}$$

The electric and magnetic fields are now easily determined with equation 9 and the relationships of equations 6 and 7.

For the method of analysis as just outlined, it is convenient to make certain assumptions which simplify the actual configuration in the z-plane in order to reduce the mathematical complexity of the Swartz-Christoffel transformation. In general, simplification takes the form

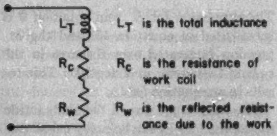

L_T
R_C
R_W
L_T is the total inductance
R_C is the resistance of work coil
R_W is the reflected resistance due to the work

Fig. 5 (left). Representation of loaded coil

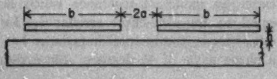


Fig. 6 (below, left). Assembly cross section

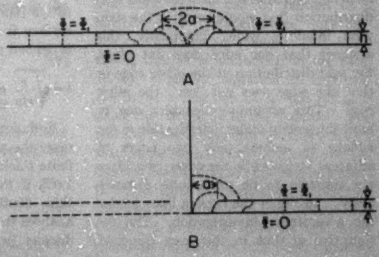


Fig. 7 (right). A—Simplified zig-zag cross section. B—Equivalent representation of A

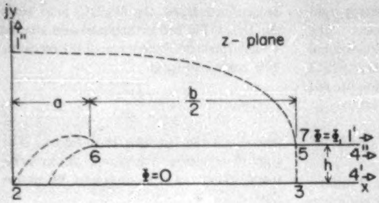


Fig. 8 (left). Cross section in z-plane

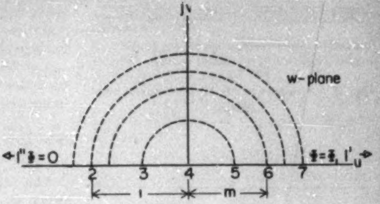


Fig. 9 (right). Cross section in w-plane

of reducing the number of "corners" in the z-plane.

First, the work coil and the work (structure) together are assumed to have zero thickness in constructing the potential analog. Fig. 3(B) represents the resulting cross section after this "zero-thickness" assumption is applied to the original cross section, represented in Fig. 3(A). This assumption is a good approximation because in the TEM wave, due to proximity effect and skin effect, the currents flow in very thin areas at the upper surface of the work and at the lower surface of the coil strips. It should be mentioned that after the field distribution has been described at the surface of the conductors with the use of the zero-thickness assumption in the analog, power calculations for the actual TEM wave are based on the assumption that all conductors are at least three skin depths thick. This is obviously necessary since power calculations based on zero-thickness conductors would be meaningless. These two seemingly inconsistent assumptions introduce little error since the field distribution based on zero-thickness conductors is approximately the same as the distribution based on small finite-thickness conductors. This is especially true for the flat-type, close-coupled coils.

The next simplification is based on the width of the work-coil strips being large compared to the spacing between the coil and the work (insulation spacing). Since this is true, the fields at the underside of the strip away from the edges will be uniform and therefore will have the same field distribution that would result if the strip were infinitely wide. Practically, this assumes that one edge does not affect the field distribution at the other edge or that one edge does not "see" the other edge. This assumption enables one to limit the section under investigation in the z-plane to include only two edges of adjacent strips while the other two edges are assumed to lie an infinite distance away. In Fig. 4 the broken lines border the area under investigation. The assumption is that in this area the field distribution in (B) is the same as the distribution in (A). The fields in other areas of (A) follow from symmetry.

Therefore, using the simplifying assumptions in the foregoing discussion, Fig. 4(B) represents the configuration in the z-plane. After the field distribution has been described, the various characteristics of the work-coil circuit can be determined.

Unless otherwise indicated, all expressions used are in the meter-kilogram-second practical system of units. In this system, the permeability of free space is $4\pi\times10^{-7}$ henrys per meter.

Power Calculations

If a sinusoidally time-varying magnetic field of rms value $H(x)$ exists at the surface of a plane conductor of infinite thickness, the time-average power density at any point on the surface is given by

$$P=H(x)^2R_s \text{ watts per meter}^2 \qquad (10)$$

where R_s is the surface resistivity defined as

$$R_s=\sqrt{\pi f\mu\rho} \text{ ohms} \qquad (11)$$

where

f = time-varying frequency, cycles per second

μ = permeability of conductor, henrys per meter

ρ = electrical resistivity of conductor, ohm-meters

The magnitude of the current in the conductor decreases exponentially with penetration into the conductor. The depth at which the current has decreased to $1/e$ (about 36.9%) of its value at the surface is termed one "skin depth" and is defined as

$$\delta=\sqrt{\frac{\rho}{\pi f\mu}} \text{ meters} \qquad (12)$$

Equation 10 can be used to approximate the power density in conductors of finite thickness with an error of less than 1.0% if the conductor is at least three skin depths thick. In the following analysis it will be assumed that the conductors (work and work-coil) meet this criterion, and equation 10 will be used in power calculations.

After equation 9 has been solved (it describes the magnetic field $H(z)$ everywhere in the z-plane), the expression for $H(z)$ along the conductor surface is substituted into equation 10 and the resulting equation integrated over the area under investigation. The resulting expression for the total power is then divided by the total area which gives the average power density over the area. This procedure can be applied to the work surface and the work-coil surface and will give the power coupled, the power loss, and the efficiency.

Impedance Calculations

The equivalent impedance of the loaded work coil may be represented by the series circuit shown in Fig. 5. Expressions for the resistive components R_c and ${}_w$ are obtained by dividing the power expressions by the square of the current.

The total inductance L_T may be divided into two parts: the inductance due to fields outside the conductors, referred to as the external inductance L_e, and the inductance due to the fields inside the conductors (work and work coil), referred to as the internal inductance L_i.

The external inductance L_e is calculated with the use of the equation which relates the instantaneous current in the conductor to the instantaneous stored energy in the magnetic field. This equation is[1]

$$\frac{1}{2}LI^2=\int_v\frac{\mu|H|^2}{2}\,dv \qquad (13)$$

The expression for H from equation 9 is substituted in equation 13, and the expression integrated over the area in the z-plane (assuming unit depth). This results in an equation for L_e.

The inductance due to the fields inside the conductors L_i can be easily calculated on the basis of semi-infinite conductors. For a semi-infinite plane conductor there exists an inductive reactance which is equal to the skin-effect resistance. In other words, the internal impedance has a phase angle of 45 degrees. Since the

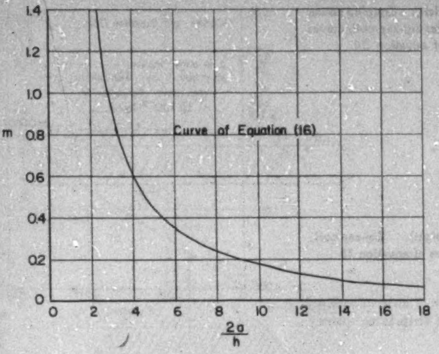


Fig. 10 (above). Curve of equation 16

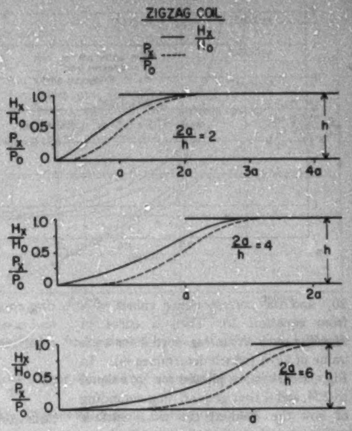


Fig. 11 (right). Field and power profile for zig-zag coil

work coil is assumed to be approximately a semi-infinite conductor (at least three skin depths thick) its internal inductive reactance will be equal in magnitude to the resistance of the work coil. Also it can be shown that, assuming the work is a semi-infinite plane conductor, the inductive reactance due to the fields in the work is equal in magnitude to the reflected resistance due to the work. Therefore, the magnitude of the reactance due to L_t is equal to the resistance "looking into" the loaded work coil.

The impedance of the loaded circuit is now fully defined. For many purposes it is convenient to define a quantity Q as

$$Q=\frac{\omega L_T}{R_c+R_w} \tag{14}$$

In the following analysis an equation for Q is derived in terms of the work-coil dimensions. Although the author has analyzed both types of work coils (zig-zag and spiral), the analysis of only the zig-zag coil is included in this paper.

Zig-Zag Work Coils

CONSTRUCTION OF POTENTIAL ANALOG

Fig. 6 represents the cross section of the work-coil assembly and shows the notation which will be used to indicate the dimensions of the configuration. In the simplified cross section of the zig-zag coil, Fig. 7(A), the conductors are given appropriate potentials Φ to result in the desired field distribution. The work-coil strips are given equal and opposite potentials and the work is at zero potential. The dashed lines indicate probable electric-flux lines.

To make the problem even simpler consider the case of one work-coil conductor in a zero-potential corner; see Fig. 7(B). The field distributions in the upper-right-half planes of Fig. 7(A) and (B) are the same because of the image effect produced by the corner in Fig. 7(B). The field distribution in the upper-left-half plane follows from symmetry.

The configuration of Fig. 7(B) is now drawn in the z-plane (Fig. 8). Transforming the configuration by the Swartz-Christoffel transformation into the w-plane (Fig. 9) gives the transformation[3]

$$z=\frac{h}{\pi}\left[\frac{2}{m}\sqrt{w+1}+\ln\frac{\sqrt{w+1}+1}{\sqrt{w+1}-1}\right]+jh \tag{15}$$

where m is defined by the correspondence of point 6 in both planes and is given by

$$\frac{2a}{h}=\frac{2}{\pi}\left[\frac{2}{m}\sqrt{m+1}+\ln\frac{\sqrt{m+1}+1}{\sqrt{m+1}-1}\right] \tag{16}$$

Equation 16 is plotted in Fig. 10.

Note that the numbers in Figs. 8 and 9 represent points which correspond in both planes. Also corresponding flux lines are represented in the two planes.

The complex potential function in the w-plane is by observation:[3]

$$P(w)=\frac{-j}{\pi}\Phi_1\ln w \tag{17}$$

MAGNETIC-FIELD DISTRIBUTION

Differentiating equations 15 and 17 and substituting in equation 8 gives for the magnetic field

$$H(z)=\frac{\Phi_1}{\eta h}\,\frac{m\sqrt{w+1}}{(w-m)} \tag{18}$$

Although the magnitude of Φ_1 is purely arbitrary, it can be shown that $\Phi_1/\eta h$ is the magnitude of the constant (space-wise) magnetic field which appears between the work-coil and the work as x approaches infinity in the z-plane. Practically, however, this is the value of the field in the region of points 3 and 5. For convenience let

$$H_0=\frac{\Phi_1}{\eta h} \tag{19}$$

where H_0 is the magnitude of this constant magnetic field.

It is of interest to examine the fields at the surface of the work. In the w-plane, the work surface (point 2 to point 4, (Fig. 9) is along the u-axis from u equal to negative unity to u equal to zero. Therefore an expression for the surface field H_z can be obtained from equation 18 by letting w equal u and using the identity of equation 19. Equation 20 gives the ratio of H_z at any point to the maximum field H_0 as x approaches infinity.

$$\frac{H_z}{H_0}=\frac{m\sqrt{u+1}}{(u-m)} \tag{20}$$

The ratio H_z/H_0 can be plotted as a function of x (distance along work surface) with m as a parameter. (Fig. 9 has shown that m is a function of the ratio $2a/h$.) However, equation 20 expresses H_z/H_0 as a function of u, the distance along the conductor in the w-plane. The transformation equation 15 relates x to u since at the work surface $z=x$ and $w=-u$, but it is impractical to express u as a function of x from this equation. Therefore the best procedure is to pick values of u (from -1 to 0) and solve for the corresponding values of H_z/H_0 from equation

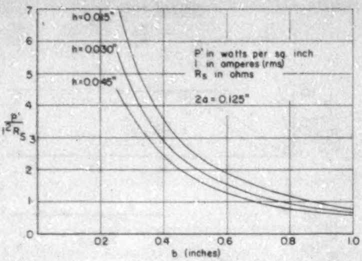


Fig. 12 (left). Coupled power density for zig-zag coil; curves of equation 34

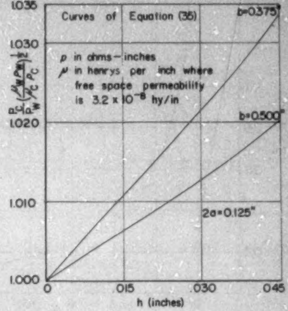


Fig. 13 (right). Zig-zag coil; curves of equation 35

20, and the corresponding values of x from equation 15. Then a curve of H_x/H_0 versus x can be plotted for each value of $2a/h$ (which determines m). In Fig. 11 curves are plotted for $2a/h$ equal to 2, 4, and 6 that give an understanding of how the field at the work surface varies as a function of x and $2a/h$.

POWER CALCULATION

As indicated in equation 10, the power varies as the square of the magnetic field. Therefore, assuming constant surface resistivity, the ratio P_x/P_0 (the ratio of the power at any point on the surface to the constant power as x approaches infinity) varies as $(H_x/H_0)^2$. These curves also appear in Fig. 11.

To make calculations of the actual magnitude of power, the value of H_0 must be expressed in terms of the coil dimensions and the current flowing in the conductors. Consider the well-known equation

$$\oint H\cdot dl = I \tag{21}$$

which states that around any closed path the integral of the dot product of the magnetic-field intensity and the elemental distance at each point on the path is equal to the current enclosed inside the path. Now, consider half of a work-coil strip as represented in Fig. 8. Point 7 represents a point on the top surface of the strip at $x=a+b/2$, and 5 is a point on the lower surface at $x=a+b/2$. A closed path along the surface of this half strip would be followed by going from point 5 to point 6 and then to point 7 since the conductor has zero thickness. This same path in the w-plane is the straight line from point 5 to point 7. For this half-strip, equation 21 becomes

$$\int_5^7 H x dx = I_1 \tag{22}$$

where I_1 is the current in the half-strip. From equation 20

$$H_x = H_0 \frac{m\sqrt{u+1}}{(u-m)} \tag{23}$$

Differentiating equation 15, while letting $z=x+jh$ (position of strip in z-plane) and $w=u$ (position of strip in w-plane), results in

$$dx = \frac{h}{\pi m}\,\frac{(u-m)du}{u\sqrt{u+1}} \tag{24}$$

Substituting equations 23 and 24 into equation 22 gives

$$I_1 = \frac{H_0 h}{\pi}\int_5^7 \frac{du}{u} \tag{25}$$

Performing the integration in 25 results in

$$I_1 = \frac{H_0 h}{\pi}\left(\ln|u_7| - \ln|u_5|\right) \tag{26}$$

where $\ln|u_5|$ is the natural logarithm of the absolute value of u at point 5 and $\ln|u_7|$ is the natural logarithm of the absolute value of u at point 7.

Now, u_5 and u_7 must be expressed in terms of the configuration dimensions a and h. From correspondence of points 5 and 7 in the two planes as expressed in the transform of equation 15, u_5 and u_7 both must satisfy the equation

$$a+b/2 = \frac{h}{\pi}\left(\frac{2}{m}\sqrt{u+1} + \ln\left|\frac{\sqrt{u+1}+1}{\sqrt{u+1}-1}\right|\right) \tag{27}$$

For $b>>h$, u_5 is much smaller than unity so that

$$\sqrt{u_5+1}\approx 1 \tag{28}$$

Therefore, when expressing u_5, equation 27 may be written as

$$\ln\left|\frac{\sqrt{u_5+1}+1}{\sqrt{u_5+1}-1}\right| = \frac{\pi(a+b/2)}{h} - \frac{2}{m} \quad (b>>h) \tag{29}$$

Since u_5 is much smaller than unity, equation 29 can be simplified (by using the first two terms of the series expansion of $(u5+1)^{1/2}$) to

$$-\ln|u_5| = \frac{\pi(a+b/2)}{h} - \frac{2}{m} - \ln 4 \tag{30}$$

It can be shown that when $b\geqq 6h$ and $2a \leqq 0.6h$, the logarithm term in equation 27 is insignificant when expressing u_7. With these restrictions $\ln|u_7|$ is approximated by

$$\ln|u_7| = \ln\left\{\left[\frac{\pi m}{2h}(a+b/2)\right]^2 - 1\right\} \tag{31}$$

By trial it was discovered that the sum of equations 30 and 31 does not have a restriction on the value of $2a/h$ for a good approximation of $\ln|u_7| - \ln|u_5|$. Although there is a restriction on equation 31 of $2a<0.6h$, $\ln|u_7|$ becomes much smaller than $-\ln|u_5|$ at values of $2a/h$ where the error in equation 31 is significant. Therefore there is very little percent error in the sum when the approximations are used. With only the restriction $b\geqq 6h$, the sum of equations 30 and 31

$$\ln|u_7| - \ln|u_5| = \ln\left\{\left[\frac{\pi m}{h}(a+b/2)\right]^2 - 1\right\} + \frac{\pi(a+b/2)}{h} - \frac{2}{m} - \ln 4 \quad (b\geqq 6h) \tag{32}$$

gives a good approximation of $\ln|u_7| - \ln|u_5|$ (equation 32).

Equation 26 may be rewritten

$$H_0 = \frac{I\pi}{2h}\,\frac{1}{(\ln|u_7| - \ln|u_5|)} \tag{33}$$

where $I=2I_1$, the total strip current, and $\ln|u_7| - \ln|u_5|$ is determined from equation 32. When equation 32 is used, it should be kept in mind that the approximations of equations 30 and 31 must be valid. If the approximations are not valid, u_5 and u_7 must be obtained from equation 27.

The average power density developed in the work is obtained by substituting equation 23 into equation 10, integrating equation 10 over the area from point 2 to point 3 (assuming unit conduction length), and then dividing by the area between points 2 and 3. The resulting equation

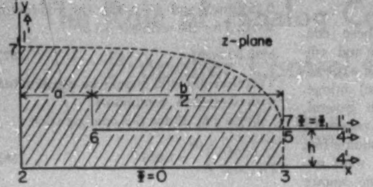


Fig. 14 (above, left). Area occupied by magnetic field from half-strip

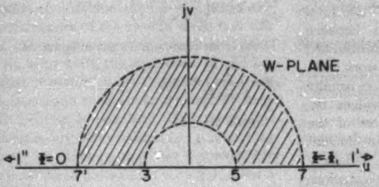


Fig. 15 (left). Area occupied by magnetic field from half-strip

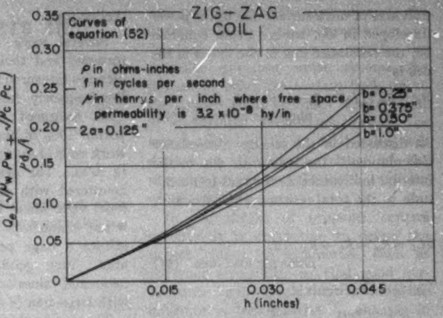


Fig. 16 (above). External Q for zig-zag coil

is as follows:

$$P'=\frac{R_sI^2\pi}{4h^2(2a+b)(\ln|u_7|-\ln|u_5|)^2}\times[\pi(2a+b)-2a\pi\sqrt{m+1}+4h]$$

watts per unit area (34)

where P' is the average power density developed in the work, R_s is the surface resistivity of the work, and I is the rms work-coil current. (Note: In deriving equation 34 it was again convenient to make the approximation of equation 28. The author considers that this approximation will be good for all practical cases.)

Fig. 12 shows the nature of equation 34. Here the width of the strip b is the independent variable with h as a parameter. The distance between strips was chosen to be 0.125 inch.

EFFICIENCY

The power lost in the work coil can be obtained in the same manner as was the power in the work by integrating from point 5 to point 7. The expressions for the power coupled in the work and that lost in the coil may express the ratio

$$\frac{P_c}{P_w}=\left(\frac{\mu_c\rho_c}{\mu_w\rho_w}\right)^{1/2}\left(1+\frac{\alpha}{\beta}\right) \tag{35}$$

where

P_c = power lost in work coil
P_w = power coupled into the work
μ_c = permeability of work-coil material
μ_w = permeability of work material
ρ_c = electrical resistivity of the work-coil material
ρ_w = electrical resistivity of work material

and

$$\alpha=4h\sqrt{m+1}\tanh^{-1}\left(\frac{4h\sqrt{m+1}}{\pi m(2a+b)}\right) \tag{36}$$

and

$$\beta=\pi(2a+b)-2a\pi\sqrt{m+1}+4h \tag{37}$$

Curves of equation 35 are plotted in Fig. 13 as a function of the spacing h with $2a=0.125$ and b as the parameter.

The efficiency of the work coil is given by

$$\text{efficiency}=\frac{P_w}{P_w+P_c}=\frac{1}{1+\dfrac{P_c}{P_w}} \tag{38}$$

and may be calculated by using the results of equation 35.

THE Q

As discussed in the section "Impedance Calculations," the external inductance is determined by the energy stored in the external magnetic field. Consider the z-plane represented in Fig. 14. All of the magnetic field produced by the half-strip is included in this area since the electric-field lines 3-5 and 7-7′ determine the outer limits of the field. In the w-plane (Fig. 15) this area lies between two semicircles.

As implied by equation 13, the external inductance is given (assuming unit conductor length) by

$$\frac{1}{2}L_eI^2=\frac{\mu}{2}\iint|H(z)|^2dxdy \tag{39}$$

Due to the conformality of mapping, the energy is also given by

$$\frac{1}{2}L_eI^2=\frac{\mu}{2}\iint|H(w)|^2dudv \tag{40}$$

Differentiating equation 17 as in equation 5 and using equation 19, the magnetic-field distribution in the w-plane is given by

$$H(w)=\frac{H_0h}{\pi}\,\frac{1}{w} \tag{41}$$

Taking the absolute value of equation 41 and squaring it gives

$$|H(w)|^2=\frac{H_0^2h^2}{\pi^2}\,\frac{1}{u^2+v^2} \tag{42}$$

Substituting equation 42 into equation 40 and changing to polar co-ordinates results in (noting that total energy due to current in the entire strip is twice that due to the half-strip):

$$\frac{1}{2}L_eI^2=\frac{\mu_dH_0^2h^2}{\pi^2}\int_0^\pi\int_5^7\frac{drd\theta}{r}=\frac{u_dH_0^2h^2}{\pi}(\ln|u_7|-\ln|u_5|) \tag{43}$$

where $\ln|u_7|-\ln|u_5|$ is defined by equation 32 and μ_d is the permeability of the dielectric material. Substituting equation 33 into equation 43 and solving for L_e, and dividing by the work surface area for one strip $(2a+b)$ gives

$$L_e'=\frac{\pi\mu_d}{2(\ln|u_7|-\ln|u_5|)(2a+b)}$$

henrys per unit area (44)

where L_e' is the average external inductance per unit of work-surface area.

An expression for the resistance due to the work can be obtained from equation 34 by dividing by the square of the current, or

$$R_w'=\frac{P'}{I^2},\ \text{ohms per unit area} \tag{45}$$

where R_w' is the average resistance per unit area of work surface due to the power developed in the work. From equation 38, the resistance due to the losses in the coil is

$$R_c'=R_w'(1/\text{eff (efficiency)}-1),\quad \text{ohms per unit area} \qquad (46)$$

As discussed in the section "Impedance Calculations," the reactance due to the internal inductance L_i is equal in magnitude to the total resistance of the circuit, or

$$\omega L_i'=R_w'+R_c'=R_w'/\text{eff},\quad \text{ohms per unit area} \qquad (47)$$

The Q of the circuit is

$$Q=\frac{\omega(L_e'+L_i')}{R_w'+R_c'}=\text{eff}\ \frac{\omega L_e'}{R_w'}+1 \qquad (48)$$

From equations 44, 45, 34, 35, and 38

$$Q=\left[\frac{4h^2\mu_d\sqrt{\pi f}(\ln|u_7|-\ln|u_8|)}{\beta(\sqrt{\mu_w\rho_w}+\sqrt{\mu_c\rho_c})+\alpha\sqrt{\mu_c\rho_c}}\right]+1 \qquad (49)$$

Since in most cases (see Fig. 13) β is much larger than α, Q can be approximated as

$$Q=\left[\frac{4h^2\sqrt{\pi f}\mu_d(\ln|u_7|-\ln|u_8|)}{\beta(\sqrt{\mu_w\rho_w}+\sqrt{\mu_c\rho_c})}\right]+1 \qquad (50)$$

Equation 50 can be written

$$Q=Q_e+Q_i \qquad (51)$$

where Q_i is the ratio of the internal impedance to the total resistance and is equal to unity, and Q_e is the ratio of the external reactance to the total resistance and is equal to

$$Q_e=\frac{4h^2\mu_d\sqrt{\pi f}\,(\ln|u_7|-\ln|u_8|)}{\beta(\sqrt{\mu_w\rho_w}+\sqrt{\mu_c\rho_c})} \qquad (52)$$

Equation 51 is plotted in Fig. 16 as a function of h with $2a=0.125$ and b as the parameter.

Application

Configuration Requirements

The equations derived in this analysis can be applied to any flat-type, close-coupled, zig-zag work coil whose configuration and application fit the description in the previous discussion. The most important configuration requirement is

$$b\geq 6h \qquad (53)$$

In general, strips that satisfy equation 53 form a wide, close-coupled work coil which validates the wide-strip assumption and satisfies the requirements for the approximations in the analysis.

Accuracy

Several work coils have been constructed and their resistances and efficiencies measured with the use of a special low-frequency (around 300 kc) bridge. This measured data was, in general, within 10% of the calculated values for work coils of approximately 2 square feet in area. The greatest errors were encountered with small-area (one square foot) coils where the end effects (places where conductors change direction) constitute a large percentage of the total area. Very good correlation between measured values and theory was obtained with large-area (4 square feet) work coils.

The measured inductances were usually greater than the calculated values because of the added inductances of the leads. At present, there are insufficient experimental data to determine definitely the accuracy of the analytic expressions for Q.

Sample Calculation

For an example, consider the case where 2 square feet of magnesium alloy is to be heated with a power density of 500 watts per square inch with a generator used which is capable of producing 800 amperes and operating at a frequency of 400 kc. It will be assumed that the work material is at least three skin depths thick. A copper, zig-zag work coil will be used.

The following information is given:

$\rho_w=3.5\times10^{-6}$ ohm-inches (magnesium alloy)
$\rho_c=6.7\times10^{-7}$ ohm-inches (copper)
$\mu_w=\mu_c=\mu_d=3.2\times10^{-8}$ henrys per inch, meter-kilogram-seconds, nonmagnetic (54)

First, a value for the spacing between conductors ($2a$) is chosen. Past experience has indicated that $2a=1/8$ inch (0.125 inch) is a good choice for most applications. Also, the necessary insulation thickness h is estimated considering the probable voltage across the coil and dielectric strength of the insulation. In this example, $h=0.030$ inch is used.

The surface resistivity of this magnesium alloy at 400 kc is

$$R_s=(\pi f\mu\rho)^{1/2}=[\pi(4\times10^5)(3.2\times10^{-8})\times(3.5\times10^{-6})]^{1/2} \qquad (55)$$
$$=3.7\times10^{-4}\text{ ohms}$$

In Fig. 12, the ordinate is

$$\frac{P'}{I^2R_s}=\frac{500}{(800)^2(3.7\times10^{-4})}=2.1 \qquad (56)$$

From Fig. 12, using the curve for $2a=0.125$ inch and $h=0.030$, the conductor width b for the ordinate 2.1 is

$$b\approx0.50\text{ inch} \qquad (57)$$

Therefore the coil should be constructed with 0.50-inch conductors.

From Fig. 13 for this configuration

$$\frac{P_c}{P_w}\left(\frac{\mu_w\rho_w}{\mu_c\rho_c}\right)^{1/2}=1.013 \qquad (58)$$

Using the given values from equation 54

$$\frac{P_c}{P_w}=1.013\left(\frac{6.7\times10^{-7}}{3.5\times10^{-6}}\right)^{1/2}=0.44 \qquad (59)$$

Then, from equation 38 the efficiency is

$$\text{eff}=\frac{1}{1+0.44}=69\% \qquad (60)$$

From Fig. 16

$$\frac{Q_e(\sqrt{\mu_w\rho_w}+\sqrt{\mu_c\rho_c})}{\mu_d\sqrt{f}}=0.13 \qquad (61)$$

Substituting the values from equation 54 into equation 61 gives

$$Q_e=5.5 \qquad (62)$$

From equation 51, the total Q is

$$Q=5.5+1=6.5 \qquad (63)$$

With this information, the characteristics of the work coil are completely defined. The reflected resistance due to the work is given by

$$R_w=\frac{\text{(power density)(area)}}{\text{current}^2}=0.23\text{ ohm} \qquad (64)$$

The total resistance R_T is

$$R_T=\frac{R_w}{\text{eff}}=\frac{0.23}{0.69}=0.33\text{ ohm} \qquad (65)$$

The total inductive reactance is

$$\omega L_T=R_TQ=0.33(6.5)=2.1\text{ ohms}$$

The total inductance is

$$L_T=\frac{\omega L_T}{\omega}=\frac{2.1}{2\pi(4\times10^5)}=0.84\text{ millihenrys} \qquad (66)$$

The total voltage across the work coil is

$$V=I|Z|=IR_T(1+Q^2)^{1/2}$$
$$=800(0.33)[1+(6.5)^2]^{1/2} \qquad (67)$$
$$=1{,}700\text{ volts}$$

References

1. Fields and Waves in Modern Radio (book), S. Ramo, J. R. Whinnery. John Wiley & Sons, Inc., New York, N. Y., 1944.

2. Simplified Theory of Microstrip Transmission Systems, F. Assadourian, E. Rimai. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, Dec. 1952.

3. Electromagnetic Fields (book), Ernst Weber. John Wiley & Sons, Inc., New York, N. Y., 1950.

The Role of Analog Computers in Propeller Control Design

ROY K. FRICK
NONMEMBER AIEE

IN THE PROCESS of designing and developing a dynamic system, provision should be made for analysis of the performance of the system while it is still in the design stage. One method of analysis is to write the equations of motion of each moving part, then combine these equations in such a manner that a means of solution is possible. This method is satisfactory if the system has few moving parts and thereby does not require many equations to describe its behavior. For a highly complex system this will prove to be very laborious by longhand methods and should be submitted to machine computation for solution. The motion of the parts of the system may be such that an equation describing the behavior of any one part may be difficult to write unless approximations are made. They may have to be directly simulated by a device which has the proper dynamic characteristics built into it.

The analog computer is especially suited to the analysis of control problems because it can compute the solution of differential equations and also simulate nonlinear phenomena directly. There are many forms of analog computers, some of which work by mechanical principles and others which work on electrical principles. The type of computer in widest use is that called the d-c electronic differential analyzer. It combines all the good features of the other types, such as flexibility, compactness, accuracy, and ease of maintenance.

The computer circuitry is wired in such a manner that the circuit equations are in the same form as the equations of the system under study. The voltages at various points in the circuit are analogous to the motions or positions of the parts of the system being studied. Because of these analogous voltage operations, the machine is called an analog computer. It is also called a simulator, although this term is generally applied to a special-purpose analog computer.[1]

Nomenclature

I = moment of inertia
N = rotational speed, rpm
ΔN = change in speed
β = propeller blade angle
$\Delta\beta$ = change in blade angle
ω_f = engine fuel flow
$\Delta\omega_f$ = change in fuel flow
Q_P = propeller load torque
ΔQ_P = change in load torque
Q_e = engine torque
ΔQ_e = change in engine torque
PLA = power lever angle
ΔPLA = change in PLA
T_4 = turbine inlet temperature
τ = time constant
(s) = Laplace operator
e_i = amplifier input voltage
e_0 = amplifier output voltage
Z_f = feedback impedance
Z_i = input impedance
i = instantaneous current
v = instantaneous voltage
X_c = capacitive reactance
J = advance ratio
V = air speed

Propeller Control Problem

A propeller is a device for converting mechanical shaft power into a propulsive force called "thrust." Thrust is a function of many variables, among these being the pitch of the propeller blades, the rotational speed of the propeller, and the air density. There are other variables, but these depend on the particular propeller being used. With the variable pitch propeller, a control system must be devised so that proper operating blade angle may be maintained. The type of control system most widely used governs the rotational speed of the engine by varying the propeller blade angle. By doing this, the control matches the propeller to the power of the engine.

With such a control system, the propeller governor must be able to sense an off-speed condition and then translate this into terms of blade angle. If an overspeed condition exists, the governor must increase the blade angle, which in turn will cause the load of the propeller on the engine to increase. This increase in load will cause the system to decelerate until such time as the rpm is again down to the desired value. If an overcorrection, or "undershoot" in this case, results, the propeller governor must reverse the process to bring the system back up to the desired value of rpm. The amount of succeeding corrections before the rpm settles out to its final value will depend on the characteristics of the propeller control and the flight condition of the aircraft.

In the design of propeller control systems, it is imperative to define the proper values of parameters, such as governor sensitivity, so that good performance will result. If a propeller governor is too sensitive to changes in rpm, excessive oscillating or "hunting" will result. If the governor and/or control system is too sluggish, excessive excursions in rpm may result. The sensitivity of the control must be a compromise such that all flight conditions of the aircraft are taken into account. For instance, a given change in propeller blade angle results in a change in propeller load torque which is larger at high air speeds than at low, and larger at low altitudes than it is at high altitudes.

The fuel control of the engine has certain dynamic properties and these must be recognized when a propeller control system is designed. Of course, the propeller control will have inherent characteristics of its own which the fuel control designer must appreciate. In any case the propeller control and the fuel control must be compatible. There must not be undesirable interactions or conditions under which the propeller control and fuel control have such unsatisfactory phase relationships that each control can oppose the corrections of the other.

The parameters which will determine proper compatibility of the propeller and engine control systems are such factors as time constants of servo systems within the propeller control, backlash in gears, true time lags of components, and dead zones of actuators.

To determine the proper values for the various parameters of the propeller control, the designer must make an analysis of the effect of each on the performance of the system. Methods of analysis generally involve approximations, but these are at a minimum if an analog computer is utilized for this purpose.

The gas turbine, or turbo-prop, engine is considered throughout this paper because this type of engine is in use in the latest propeller developments.

Paper 57-501, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, March 7–9, 1957. Manuscript submitted February 4, 1957; made available for printing March 14, 1957.

ROY K. FRICK is with the Wright Air Development Center, Dayton, Ohio.

Discussion

Analog Computer Fundamentals

To understand the principles of analog computation, one must start with the fundamental concept of the transfer function. In Fig. 1 is shown the standard symbol of a transfer function. The transfer function G is said to operate on the input quantity e_i in such a way as to produce the output e_0. The relation is such that

$$e_0 = Ge_i$$

The transfer function concept is applied to an operational amplifier (Fig. 2), the basic component of a d-c electronic analog computer. The amplifier has input and output terminals which are accessible so that impedances may be wired into the circuitry of the amplifier. If an impedance is wired across the input and output terminals, it is referred to as a feedback impedance. If an impedance is wired only on the input of the amplifier, it is referred to as the input of the amplifier. The theory of the circuit design of the amplifier is such that the transfer function of the amplifier is equal in magnitude to the ratio of the feedback and input impedances and is opposite in sign. The transfer function G of the amplifier can be written

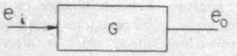


Fig. 1. Transfer function

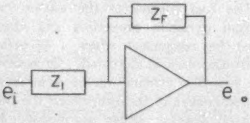


Fig. 2. Computer amplifier

$$G = -\frac{Z_f}{Z_i} = \frac{e_0}{e_i}$$

where Z_f=feedback impedance, and Z_i=input impedance.

Depending on the types of impedances, in this case capacitive or resistive, an amplifier can serve as an integrator, a differentiator, or it may multiply by a constant. By wiring in various input impedances in parallel, an amplifier may serve as a summer, i.e., the various input quantities can be added. When the feedback impedance is capacitive and the input is resistive, the amplifier is an integrator If the process is reversed, i.e., if the input is capacitive and the feedback resistive, the amplifier becomes a differentiator. By making both the feedback and input impedances resistive, the amplifier merely multiplies the input signal by the ratio of the feedback and input impedances. The theory of the transfer function of an operational amplifier is discussed in Appendix I. By utilizing the transfer function principal of the amplifier, one may proceed to set up a circuit which will solve linear differential equations with constant coefficients.

Basic Mathematics of Propeller Control System

A simplified engine and propeller system may be represented by a set of equations, some of which are differential equations. The equations given are representative of a propeller with a gas turbine engine.

The engine torque must accelerate the polar mass of the system and overcome the load of the propeller:

$$Q_e = 2\pi I \frac{d}{dt} N + Q_P \tag{1}$$

or if only increments of torque and rpm are considered and equation 1 is rearranged

$$\Delta Q_e - \Delta Q_P = 2\pi I \frac{d}{dt} \Delta N \tag{2}$$

The engine torque may be expressed as a function of fuel flow and rpm and the incremental equation for engine torque is

$$\Delta Q_e = \frac{\partial Q_e}{\partial N}\Big|_{\omega_f} \Delta N + \frac{\partial Q_e}{\partial \omega_f}\Big|_N \Delta\omega_f \tag{3}$$

Similarly the incremental propeller torque equation may be written with the knowledge that propeller torque is a function of blade angle and rpm.

$$\Delta Q_P = \frac{\partial Q_P}{\partial N}\Big|_\beta \Delta N + \frac{\partial Q_P}{\partial \beta}\Big|_N \Delta\beta \tag{4}$$

Equations 2, 3, and 4 may be combined:

$$\left(\frac{\partial Q_e}{\partial N} - \frac{\partial Q_P}{\partial N}\right)\Big|_{\beta\omega_f} \Delta N + \frac{\partial Q_e}{\partial \omega_f} \Delta\omega_f - \frac{\partial Q_P}{\partial \beta} \Delta\beta = 2\pi I \frac{d}{dt} \Delta N \tag{5}$$

The following equation is generally used to express the operation of the propeller governor:

$$\Delta\beta = A \int^t \Delta N dt + \beta \Delta N \tag{6}$$

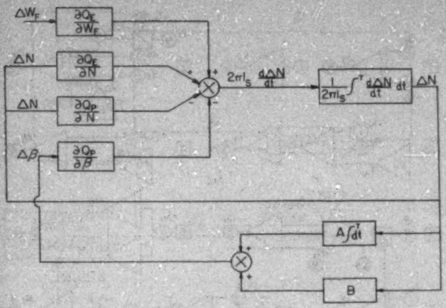
One extensively used practice is to draw a block diagram of the system before proceeding with the setup of the analog circuit. A block diagram of the system which has been described by the previous equations is presented in Fig. 3.

This block diagram is a collection of transfer functions connected together to form an over-all control loop. The advantage of the block diagram is that it shows the relationship of all the parameters and is a guide in setting up the analog circuit. The circuit after being set up is generally a duplicate of the block diagram. In Fig. 3, it is seen that each block is a transfer function and that these are determined from the coefficients of equation 5. The computer diagram for this block diagram is presented in Fig. 4. In Fig. 4 the plain triangular figures are summing amplifiers and the triangular figures with a double line are integrators. In many analog computers, the input and feedback impedances are internally wired and one has a choice of feedback and input combinations on a terminal board. When this is the case, symbols for summers and integrators are as those used in Fig. 4.

Fuel flow, rpm, and blade angle are multiplied by the appropriate partials in amplifier *1* which in turn produces a signal proportional to rate of change of rpm with a change in sign. Amplifier *2* integrates this signal and produces rpm increment with another change in sign. Rpm is integrated in amplifier *3* and changed in sign in amplifier *4*. These two outputs are added in amplifier *5* to produce positive blade angle increment. Since negative blade angle increment is needed for an input signal to amplifier *1*, amplifier *6* is needed to change the sign of the output of amplifier *5*. To complete the loop, rpm increment with the proper sign is fed from the output of amplifier *4* into the input of amplifier *1*. A check on the circuit for stability can be made by following the rule that there must be an odd number of amplifiers in any one loop. Reference to Fig. 4 indicates that this is true.

Fig. 4 shows that the fuel flow increment is the independent variable and that increments of blade angle and rpm are dependent variables. Other independent variables such as the change in rpm desired can be added to the circuit if appropriate for the particular system under study.

The computer circuit in Fig. 4 is a simulation of a propeller and engine system. It is a simplification, but one can use it to study the effects of the propeller and engine partials on the stability of the system at various flight



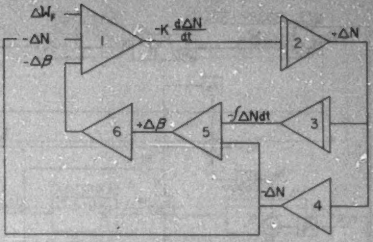


Fig. 3 (left). Linearized block diagram

Fig. 4 (above). Linearized computer circuit

conditions. Even though this circuit is not complex, the propeller designer can find satisfactory values for the partials of propeller torque to blade angle and rpm, and thereby influence blade characteristics. Satisfactory values for constants A and B can be found for equation 6 and this in turn will determine the proper governor design.

The circuit in Fig. 4, even though a simulation in probably its simplest form, is a far more effective tool in solving the equations than manpower alone. If equations 5 and 6 are combined, a third-order differential equation results and although not impossible to solve, would nevertheless consume quite a bit of effort. In addition, the solution would have to be tabulated and plotted and would not be quite as accurate as the computer solution.

FURTHER CONSIDERATIONS

The simulation and analysis which has been explained to this point is called a linearized or perturbation type of study. This is so called because the equations are linear, i.e., constant coefficients, and the variables are treated as increments about a steady-state operating point. This type of study is valuable for evaluating the system for stability under various conditions. A high degree of accuracy is possible in studying the effect of the parameters on the stability of the system.

If extending the operating range of the variables is desired, a linearized system can no longer be considered. However, it may be desired to retain a fair degree of accuracy in studying the stability of the system, but to study the effect of various nonlinearities on general system operation may not be wanted. In addition, it may be desired to determine the transient characteristics to a higher degree of accuracy than is possible from a linearized computer setup. To proceed with this type of study a set of equations may be written which are based on available engine and propeller data. The equations will be empirical and will have no standard form from engine to engine and propeller to propeller. The following equations are illustrative.

The engine and propeller torque balance will still accelerate the system.

$$Q_e - Q_P = 2\pi I \frac{d}{dt} N \qquad (7)$$

The engine torque may take the form

$$Q_e = (K_1 + K_2 T_4) N + K_3 T_4 + Q_{e_0} \qquad (8)$$

and the turbine inlet temperature

$$T_4 = K_4 N + K_5 \omega_f + T_{4_0} \qquad (9)$$

and the fuel flow

$$\omega_f = K_6 PLA + \omega_{f_0} \qquad (10)$$

and the propeller torque

$$Q_P = K_7 \beta^2 + K_8 N + Q_{P_0} \qquad (11)$$

and the propeller governor action

$$\beta = K_9 \int^t N_e dt + K_{10} N_e + \beta_0 \qquad (12)$$

The speed error is equal to the difference between the actual speed and the desired speed:

$$N_e = N - N_0 \qquad (13)$$

and the desired speed is a function of power lever angle:

$$N_D = K_{11} PLA + N_{D_0} \qquad (14)$$

The quantities with zero subscripts are initial conditions. Most makes of analog computers have special amplifiers incorporated which provide for setting and maintaining initial conditions.

The block diagram resulting from the previous equations is presented in Fig. 5. As before, the block diagram is the transfer function collection derived from a set of equations. The block diagram in Fig. 5 is similar in style to that in Fig. 3, but the circuit diagram in Fig. 6 is different in several ways from the diagram in Fig. 4. First, the diagram is much more extensive and, in an effort to simplify reading and setting up on the computer, the circuit schematic is broken up into segments. Labels such as encircled numbers are put on the various inputs and outputs; corresponding labels, in this case numbers, are joined together or are put at a common point when the circuit is being wired. One can see that if the labels were omitted on the drawing and solid lines were used for showing connections, the schematic would be harder to read and harder to follow when wiring the computer was attempted.

Second, the initial condition amplifiers are shown in the circuit. This is indicated by a triangular symbol which encloses the initials *IC*. Third, the symbol of a box with a diagonal is used to indicate a multiplier. A multiplier can multiply two or more variables or raise a variable to a power.

With this circuit, the engine and propeller system is very accurately simulated for purposes of studying transient operation over a reasonable degree of power excursions. It can also serve to study the stability to a reasonable degree of accuracy, though not as well as the linearized circuit.

Each amplifier which produces a signal in which one may be interested, say, rpm, may have the output utilized to drive a recording instrument such as an oscillograph. The various channels of the oscillograph will then record the variables as a function of time.

To simulate completely an engine-and-propeller system, the circuit in Fig. 6 would have to be further expanded. An expanded circuit is not shown, but the

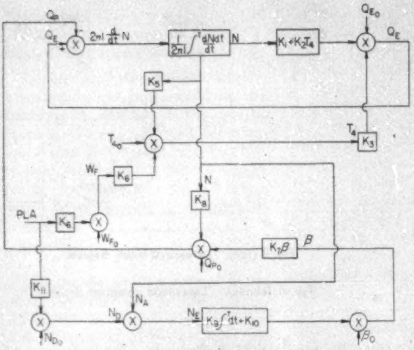


Fig. 5. Expanded block diagram

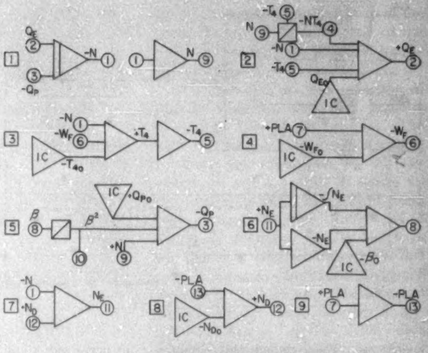


Fig. 6. Expanded computer circuit

additional equation and circuit complexities are discussed in succeeding paragraphs and in Appendix II. A lot depends on judgment as to how detailed a system should be simulated. It will depend on the factors in which one is interested and on the complexity of the system. These expanded operations will be necessary if, in the judgment of the engineer, the various second-order effects and nonlinearities are to be considered in evaluating the turboprop system performance.

ADDITIONAL COMPLEXITIES IN PROPELLER

Equation 6 describes the propeller governor action which selects the proper blade angle to maintain the desired rpm. This equation is a close approximation for some types of propeller governors over the entire operating range. For other types of propeller governors, equation 6 will be valid only for perturbation-type studies where rpm and blade angle excursions are limited to small values. One feature which practically every propeller control possesses is the use of a servo system to translate the output motion of the governor into a blade angle motion. In this case, equation 6 is actually one which describes the desired blade angle instead of the true blade angle. The relation between the desired blade angle and the actual blade angle may be written as a first-order lag with a gain G so that

$$\beta=\frac{G}{1+\tau(s)}\beta_D \tag{15}$$

This is assuming that such characteristics as dead band or backlash in gears is not significant.

In addition to this added complexity, the governor itself may employ servo systems of its own in generating the necessary terms in equation 6. Whenever these servos are employed, they may also be expressed by a first-order lag.[2,3] Equation 6 could become

$$\Delta\beta_D=\frac{A\int^t \Delta N dt}{1+\tau_2(s)}+\frac{B\Delta N}{1+\tau_3(s)} \tag{16}$$

combining equation 16 with equation 15

$$(1+\tau_1(s))\Delta\beta=\frac{A\int^t \Delta N dt}{1+\tau_2(s)}+\frac{B\Delta N}{1+\tau_3(s)} \tag{17}$$

As has been stated previously, equation 6 was valid for the entire operating range when some types of governor designs are considered. For other designs this is not true because the coefficients A and B are not constant. As an example, in many hydraulic propellers coefficients A and B are determined by the flow of fluid through a valve versus the valve displacement. The valve configuration may be a plunger which is enclosed by a sleeve with a circular hole. In equilibrium conditions, the plunger will completely cover the hole but when displaced will uncover an area of the hole so that this exposed area is a function of the square of the displacement. The flow of fluid through the valve is directly proportional to the plunger displacement. This flow may be used to operate a servo system in generating coefficient B or may determine coefficient A.

Coefficients A and B in equation 6 are slopes of the curves depicting the relation of blade angle rate to rpm error and blade angle position to rpm error respectively. When A and B are nonlinear, it is simpler to write an empirical equation directly in terms of blade angle and rpm rather than trying to determine the slopes of the curves and in turn multiplying these by rpm error.

Equation 17 may turn out to be a power series in terms of rpm error and integral of rpm error.

$$(1+\tau_1(s))\Delta\beta=\frac{\int(A_1\Delta N^3+A_2\Delta N^2+A_3\Delta N)dt}{1+\tau_2(s)}+\frac{B_1\Delta N^3+B_2\Delta N^2+B_3\Delta N}{1+\tau_3(s)} \tag{18}$$

The computer circuitry for this equation would be much more complex than that portion in Figs. 4 and 6, which simulate the governor action. If such things as dead band in governor or blade angle movement or backlash in gears are considered, the computer circuit must be further expanded to simulate these characteristics.

To raise a variable to a power such as that in equation 18, a multiplier is necessary. To generate a lag, a complex network is used for the feedback impedance of an amplifier. This feedback network consists of a capacitor and resistance in parallel. The development of this relation is left to the reader.

A circuit which can simulate dead zone is shown in Fig. 7. If the variable X is positive, but less than the absolute value of the constant A, amplifier *1* will have a positive output and the diode tube V_1 will not conduct. When the value of X exceeds the absolute value of A, amplifier *1* will have a negative

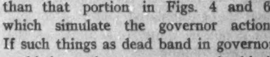


Fig. 7. Dead-band circuit

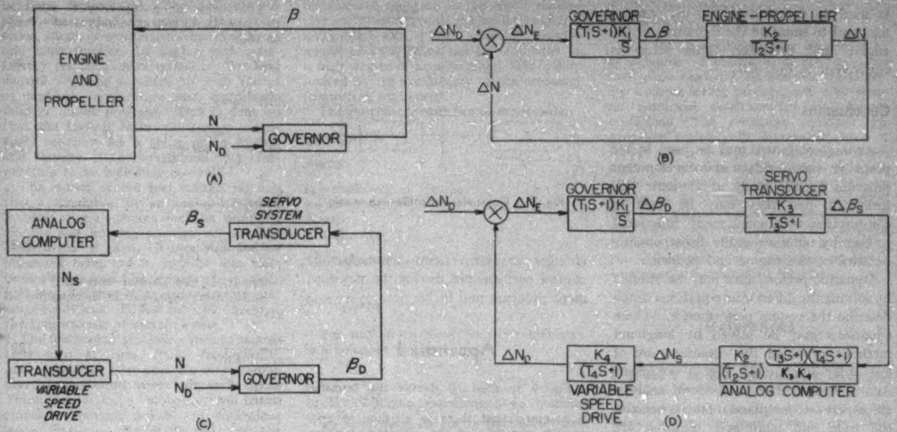


Fig. 8. Basic and equivalent system block diagrams

output and thereby allow tube V_1 to conduct. Amplifier 2 will then generate a function of X provided X exceeds A, the dead zone.

The various complexities such as lags and nonlinearities which have been mentioned can be added to the computer circuitry to simulate the system more accurately. The complexity is only limited by the equipment on hand and if the available equipment is limited, the system must be simplified. Valid approximations must be made for most problems, but made depending on the conditions of the problem. For instance, if to determine stability is desired, the equations may be linearized and characteristics such as dead band included. If it is desired to determine accurately the transient performance, empirical equations in the form of power series, may be written but such things as dead band approximated.

To evaluate properly the control performance of an engine and propeller system, a device such as an analog computer is a necessity. This is mainly because a highly nonlinear system such as an engine and propeller simply cannot be analyzed by longhand methods. Even if the equations are linearized, the analysis would possibly require the solution of a high-order differential equation which would be at best extremely long and laborious. Another method of analysis sometimes used for a system such as presented in equations 1 through 6 is called the step-by-step method. With this method values are calculated for each of the variables over a series of time increments such as 1/10 second. The initial torque unbalance between the engine and propeller is used to calculate the acceleration which in turn determines the rpm at a time interval later. The rpm is used to calculate the blade angle and in turn the engine and propeller torques. The new torque unbalance determines a new acceleration and the process is repeated. This repetition goes on until the rpm and blade angle settle out at a final value and the torque unbalance and thereby the rpm acceleration is equal to zero.

USE OF ANALOG COMPUTERS IN COMPONENT TESTING

If a set of equations can represent a system such as an engine and propeller, it can be reasoned that one or more of these equations can be replaced by the actual component or subsystem it represents. To illustrate this, reference is made to Fig. 8. The analog computer essentially takes the place of the whole system with the exception of the component under test, in this case, the governor. Stated another way, the whole system is being simulated except the governor. If the computer works in terms of voltages, suitable transducers must be provided to translate the actual motions of the governor into voltage signals to the computer and to translate the voltage signals of the computer into suitable inputs to the governor. Most transducing systems have at least a first-order lag, so care must be exercised so that compensation is made for these lags in the over-all control loop. Fig. 8(A) represents in a general way the operation of a propeller control system. The governor accepts a signal which represents rpm and compares this to the desired rpm. The governor then translates this into an output which is proportional to the blade angle which is necessary to maintain the proper rpm. Whenever the blade angle changes, the load on the engine is changed which causes a change in rpm. Fig. 8(B) is a block diagram of an engine-propeller control system. The block representing the engine and propeller show that this part of the system acts as a time lag in relation to the blade angle and the rpm. This is actually a simplification, but for the purpose of this discussion, the assumption is valid. To provide good control performance, i.e., good transient and steady-state characteristics, the governor compensates for this lag by putting a lead term into the system.[3,4] Also to provide isochronous operation, the governor possesses an integrating term. The block representing the governor is equation 6 in the form of a transfer function. In Fig. 8(C), an attempt is made to build up a configuration which is dynamically similar to (A). As has been stated before, transducers introduce lags and compensations must be made for these lags. Fig. 8(D) shows the over-all loop where an analog computer is used to compensate for the various lags and to simulate the dynamics of the engine and propeller. This setup is equivalent to Fig. 8(B).

Conclusions

Analog computers may be used in the propeller control design and development program for the study of dynamic performance, and they may be used in conjunction with component functional testing by simulating the flight characteristics of the engine and propeller.

Dynamic performance may be studied by solving the differential equations which describe the system performance. These equations may be solved by longhand methods provided the equations are of low order and are not highly nonlinear. Analog computation methods of analysis are superior to longhand methods because the time and manpower required are less and because fewer approximations are necessary. Solution of the equations results in establishment of criteria for satisfactory transient and steady-state control performance under all expected flight conditions.

The number of necessary differential and empirical equations in describing the system is determined by the number of unknown variables. Each of the equations describe the behavior of a particular subsystem or component. If one of these equations is deleted, a physical article which is described by this equation may be put into the simulation provided the proper transducers are available. When this is done, the overall system dynamics remain essentially the same, and control components can be subjected to simulated flight conditions. The analog computer can simulate the characteristics of the airplane, the engine, and the propeller blades, as well as the rest of the control system not under test.

A proper propeller control design will be possible only if adequate analysis and testing are accomplished. Analysis by use of analog computers results in the establishment of the proper values for all of the important parameters without having to resort to large-scale approximations, guesswork, or trial-and-error methods. Testing of components under simulated flight conditions will result in better ultimate performance and higher reliability.

Judicious use of analog computers in the design and development of propeller control system yields results in reduction of costs. These costs may be incurred from repeated engineering and design changes resulting from unsatisfactory control performance during the development program and in service.

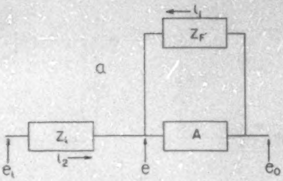


Fig. 9. Computer amplifier schematic

Appendix I

Fig. 9 is used to derive the transfer function of a computer amplifier. If it is assumed that there is no flow of grid current, the sum of the currents at point is zero, and it consists of the currents through the feedback and input impedances so that

$$i_1+i_2=0 \tag{19}$$

$$\text{but } i_2=\frac{e_i-e}{Z_2} \tag{20}$$

$$\text{and } i_1=\frac{e_0-e}{Z_1} \tag{21}$$

The output voltage is equal to the gain A of the amplifier times the grid voltage e.

$$e_0=A_e$$

$$\text{and } e=\frac{e_0}{A} \tag{22}$$

combining (20), (21), and (22)

$$\left(e_0-\frac{e_0}{A}\right)\frac{1}{Z_1}=\left(e_i-\frac{e_0}{A}\right)\frac{1}{Z_2} \tag{23}$$

and

$$\frac{e_0}{Z_1}-\frac{1}{A}\left(\frac{e_0}{Z_1}+\frac{e_0}{Z_2}\right)=-\frac{e_i}{Z_2} \tag{24}$$

The gain A is generally a very large number on the order of 10^6 so that to an approximation

$$\frac{e_0}{Z_1}=-\frac{e_i}{Z_2}$$

or

$$\frac{e_0}{e_i}=-\frac{Z_1}{Z_2} \tag{25}$$

If the impedances are pure resistances, the transfer function of the amplifier is merely the ratio of the feedback and input resistances. If one of the impedances is a reactance and the other is a resistance, the amplifier will serve as an integrator or differentiator, depending on the combination. To illustrate this, consider the case of the feedback impedance of an amplifier being equal to a capacitive reactance and the input impedance equal to a resistance.

The charge on a capacitor is equal to the capacitance times the impressed voltage so that

$$q=Cv \tag{26}$$

$$\text{but } \frac{d}{dt}q=i$$

$$\text{or } v=\frac{1}{C}\int^t i\,dt=\frac{1}{C(s)}\,i \tag{27}$$

$$\text{But } \frac{V}{i}=X_c$$

$$\text{so that } X_c=\frac{1}{C(s)} \tag{28}$$

where (s) is the Laplace operator.[2,3]

Substituting equation 28 in equation 25 gives

$$\frac{e_0}{e_i}=-\frac{1}{RC(s)} \tag{29}$$

where R is the input impedance or, restated

$$e_0=-\frac{1}{RC}\int^t e_i\,dt \tag{30}$$

Appendix II

The engine torque expressed by equations 3 and 8 is satisfactory for the intended purpose of these equations. However, in them the assumption is made that the engine torque which is delivered to the propeller shaft is instantaneous in response to the variables expressed. In other words, it is assumed that the engine torque follows the same relation to time as the fuel flow or rpm or turbine inlet temperature. For most propeller studies, it has been found that this assumption is valid and that a proper propeller control design will allow for variations in the engine torque dynamics. However, if one desires to simulate the engine more accurately, he must consider the dynamics of all the stations such as inlet duct, gear box, compressor, burner, turbine, and nozzle. Such things as pressures, air flows, temperatures, and specific heats would have to be considered and equations would have to be written to describe the various parameters in terms of known variables. A complete set of equations describing the thermal cycle of a turboprop engine is beyond the scope of this paper. One may consult any standard reference on gas turbine power plants to appreciate the complexity involved.[4,5] For the study of propeller control systems, it has been found that higher-order effects of the engine station are relatively insignificant. The dynamics of the fuel control should be considered, however, as experience has shown that interaction effects of the fuel control and propeller are very significant.

The fuel control of a typical turboprop engine is designed to maintain the fuel flow so that a nearly constant turbine inlet temperature exists for all flight conditions. This may be accomplished by a direct turbine inlet temperature controller or may be synthesized by sampling other quantities which the turbine inlet temperature depends on. The fuel control must also protect the engine from overspeed, surge, or overtemperature conditions and must provide the engine with satisfactory response characteristics. The fuel control is made sensitive to such things as ambient temperature and compressor inlet or outlet pressure. This is done so that the fuel-air ratio will remain essentially constant for a given power setting and thereby cause the turbine inlet temperature to be indirectly controlled.

The nature of the fuel control is such that a simulation by an analog computer involves special devices such as switching circuits and limiting circuits. A detailed explanation of these circuits will not be presented here, but it may be seen that these circuits would be necessary to simulate fuel flow limiting at overtemperature and surge conditions, and fuel cutback whenever the rpm exceeds at certain value.[1]

The fuel control in sensing various parameters such as rpm, PLA, temperature, and compressor inlet pressure may employ servo loops to achieve a certain output. These servo loops probably will have transfer functions which are equivalent to a first-order lag.[2,3] If each of the parameters is characterized by a lag term where significant, the equations describing the fuel flow become very complex. From these considerations, it is seen that a true simulation of the fuel control is very complex and goes considerably beyond the circuit presented in Fig. 6. However, the circuit can be made highly representative of the actual fuel control operation, provided the necessary equipment to do the job is available.

The propeller torque relationship, if simulated accurately, can be the most complicated equation in the whole simulation and thereby result in circuitry which is quite elaborate. The propeller aerodynamic characteristics can be simulated to a high degree of accuracy and this is only limited by the availability of the necessary computing equipment.

The propeller torque can be expressed as

$$Q_P = c_p \rho N^2 D^5$$

where

c_p = coefficient of torque
ρ = air density
N = rpm
D = diameter

This equation can be rewritten as

$$c_p = \frac{Q}{\rho N^2 D^5}$$

For any given propeller, the diameter is a constant so that

$$c_p = K \frac{Q_P}{\rho N^2}$$

To determine c_p for a particular propeller, at a given set of conditions, the propeller may be calibrated for torque in terms of rpm and blade angle. In practice, the torque coefficient is determined for various blade angles, diameters, rpm, and forward airspeeds by using dynamically similar propeller models in wind tunnels. The torque coefficient is generally expressed as a function of blade angle and a term called "advance ratio."[4,6]

$$c_p = f(\beta, J)$$

where the symbols represent the following:

$J = V/ND$ = advance ratio
V = forward air speed

The torque coefficient may be presented as a family of curves with advance ratio and blade angle as independent variables. By a curve-fitting process, one may write an empirical equation for the torque coefficient. The resultant equation for propeller torque will involve at least a second power term in J, the advance ratio, and may involve higher orders of blade angle and rpm. This equation can be generated to a high degree of accuracy, but may require a considerable number of multipliers and amplifiers. A multiplier is needed every time a variable is raised to the second power or higher and whenever two terms are multiplied together.

References

1. Electronic Analogue Computers (book), G. A. and T. M. Korn. McGraw-Hill Book Company, Inc., New York, N. Y., 1952, pp. 1–21.

2. Engineering Cybernetics (book), H. S. Tsien. McGraw-Hill Book Company, Inc., 1954, pp. 7–23, 34–38, 63–66.

3. How to Specify the Performance of Closed Loop System, John E. Gibson. *Control Engineering*, New York, N. Y., Sept. 1956, pp. 122–29.

4. Aircraft Jet Powerplants (book), Franklin P. Durham. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1951, pp. 47–74, pp. 88–101.

5. A Dynamic Performance Computer for Gas Turbine Engines, V. L. Larrowe, M. M. Spencer. University of Michigan, Ann Arbor, Mich., Aug. 1955, pp. 6–45.

6. Fluid Mechanics (book), R. L. Daugherty, A. C. Ingersoll. McGraw-Hill Book Company, Inc., 1954, pp. 325–31.

◆

Discussion

Irving A. Peltier (Wright Air Development Center, Dayton, Ohio): The author apparently sacrifices mathematical rigor for simplicity of presentation. Some examples follow:

He defines a transfer function G as the ratio between the output and input voltages of a system without defining these voltages adequately. If $G = e_0/e_i$ and e_0 and e_i are functions of time, then G is defined in the real time domain and its form is not particularly useful if the system contains time constants. A common method of expressing transfer functions in algebraic form is to assume that both the input and output voltages expressed as functions of time are Laplace transformable and then $G(s) = E_0(s)/E_i(s)$ where $E_0(s)$ and $E_i(s)$ are the Laplace transforms of the output and input voltages respectively. The transfer function $G(s)$ is in the complex frequency or s domain and is in a much more useful form since it describes the behavior of the output in response to many types of inputs readily.

Equation 27 is inconsistent.

$$v = \frac{1}{C}\int^t idt = \frac{1}{C(s)}i \qquad (27)$$

It should be written in two parts:

$$\text{If } v = \frac{1}{C}\int_0^t idt$$

then

$$V(s) = \frac{1}{Cs}I(s)$$

where $V(s)$ and $I(s)$ are the Laplace transforms of v and i respectively.

Equation 30 is correct if the initial value of e_0 is zero but the method of arriving at this equation is not convincing.

If the transfer function of a computer amplifier is derived with the use of Laplace transforms and the author's assumptions about grid current and output impedance, then

$$G(s) = \frac{E_0(s)}{E_i(s)} = -\frac{AZ_f(s)}{AZ_i(s) + Z_f(s) + Z_i(s)}$$

If A is very large, the approximation

$$G(s) = -\frac{Z_f(s)}{Z_i(s)}$$

is good over a large region containing values of s. However, if it is desired to differentiate e_i by making $Z_i(s) = 1/Cs$ and $Z_f(s) = R$, difficulty is encountered when e_i is changing rapidly with time. If it is desired to integrate e_i by making $Z_f(s) = 1/Cs$ and $Z_i(s) = R$, difficulty is encountered when e_i is changing slowly. The range of good approximation can be extended by proper choices of R and C but these values are very often limited by input and output impedance requirements.

The author says, "A check on the circuit for stability can be made by following the rule that there must be an odd number of amplifiers in any one loop." This neither constitutes a necessary condition nor a sufficient one for stability, even in linear systems.

It is reasonable to assume that handling nonlinear systems requires the services of an analog computer because mathematical methods are not available. However, the author should offer further proof of the analog computer's advantages over mathematical methods in the case of linear systems. It is conceivable that a mathematician might easily arrive at a single equation (transfer function) for a system. If the response of the system is too labor-consuming to plot, then he might, by modern synthesis methods, devise a passive network or semipassive network which has identically the same transfer function. This new system allows the original one to be studied for various inputs but does not allow it to be studied for changes in parameter or changes in the operation of basic components in the original system.

Economic Evaluation of Industrial Power-System Reliability

W. H. DICKINSON
MEMBER AIEE

IN THE DESIGN of an industrial plant, it is the function of the electrical engineer to determine what type of power system is required. This paper describes methods for the analysis of the relative reliability of different types of industrial power systems and the economic evaluation of the improvement in reliability which results with alternative types of systems. From this, the incremental investment that can be justified can be determined.

The paper covers four main points. First, a technique for the economic analysis of alternatives involving incremental investments is developed. The technique is a simplification of the discounted-cash-flow method of comparing alternatives. Second, probability and its application to power-system reliability problems is discussed. Third, these probability principles are applied to several typical industrial power systems. And last, a typical example is worked out. In this example, the incremental investment that is justified to increase reliability is evaluated.

Nomenclature

I = incremental investment

E = gross earnings (before income taxes, depreciation, and operating costs proportional to investment) resulting from I, expressed as uniform equivalent annual amount when nonuniform, that is,

$$E=a_n\sum_{n=1}^{n=n}\frac{E_n}{s_n}$$

R = investment charge factor

G = gross revenues or sales resulting from I

X = operating expenses (all costs not proportional to I)

C = cash flow or income (net income plus depreciation)

N = annual net income (after income taxes)

D = depreciation permitted by income tax authorities (straight line, or levelized or uniform equivalent value of sum-of-digits or declining balances), $D=j_nI$, per unit values

d_n = depreciation rate by sinking fund method, per unit value $=A_n/F_n$

$$=\frac{r}{(1+r)^n-1}=\frac{r}{s_n-1}$$

r = return rate (after taxes) per unit value, which is required on I

e = expense rate, per unit value, including all expense items proportional to I not included with X, usually includes maintenance and state taxes

t = income tax rate, per unit value

a_n = capital recovery factor $=A_n/P_1=r+d_n$, per unit value

s_n = compound return rate factor $=F_n/P_1=(1+r)^n$, per unit value

P_1 = present value of a future sum or sums of money at beginning of year one

F_n = value of a sum or sums of money in the future at the end of year n

n = number of years

A_n = annual amounts of money at the end of each year for n years, having a present value of P_1 at the beginning of the first year, and a future value of F_n at end of the nth year

$j_n=D/I$

p = probability of supply, circuit, or element not being on forced outage

f = probable annual equipment or supply failure rate causing a forced outage

S = shutdown time or forced outage hours per failure

q = probability of supply, circuit, or element being on forced outage

Technique for Economic Evaluation

In comparing power-system alternatives, the basic principle should be to determine the effect of each on the plant's cash flows. Cash flow is defined as net income plus depreciation, that is, depreciation is added to the profits after income taxes. Depreciation is included with the cash receipts or flow, since it is not an out-of-pocket expenditure. The discounted-cash-flow method of comparing alternatives follows this principle. It is believed to be superior to other methods which do not.[1-3] The discounted-cash-flow method gives a more realistic economic return irrespective of the time pattern of the earnings or the timing of investments. The differences between dollars today and dollars in the future are recognized.

In the discounted-cash-flow method, the return rate is found that discounts the cash flow or receipts of the proposed alternative design to zero, counting all investments as negative receipts. This is equivalent to saying that the present worth of all the cash receipts is equal to the present worth of all the investments. The present worth of a sum of money at the end of year n is the sum divided by the factor s_n. The present worth of a series of n equal sums at the end of each year is found by dividing the sum by a_n. By multiplying the present value of a sum by a_n the uniform equivalent annual sums at the end of each of n years are found. For an explanation of any symbols not explained in the text, reference should be made to the nomenclature. In it, also, some of the financial mathematics necessary to deal with present and future sums of money are given. A disadvantage of the discounted-cash-flow method is that its solution requires cut-and-try methods. In order to obtain the advantages of the discounted-cash-flow method, but to avoid the complexity of its cut-and-try solution, a simplified technique has been developed.

JUSTIFIABLE INVESTMENT

This technique uses the following equation, the development of which is given in Appendix I.

$$I=E/R \tag{1}$$

This is not Ohm's law. But it is somewhat analogous. It might be termed a financial Ohm's law. The investment, I, that can be justified is proportional to E, gross earnings; that is, the earnings before deduction of income taxes, depreciation, maintenance, state taxes, or any other expense that is proportional to investment. Similarly, I is inversely proportional to R. R might be termed an investment charge factor. It differs from a fixed charge factor in that it includes the required rate of return before income taxes.

The value of R in equation 1 is

$$R=\frac{r}{1-t}+e+\frac{d_n-j_nt}{1-t} \tag{2}$$

$$=\frac{a_n-j_nt}{1-t}+e \tag{3}$$

If, in equation 1, I and E are known, R may be determined. In the following equation, this value is substituted into equation 3. From which

$$a_n=(R-e)(1-t)+j_nt \tag{4}$$

By reference to tables, such as Table I, the return rate r may be found, if a_n is known. It also may be found by trial and error.

Paper 57-639, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted March 25, 1957; made available for printing April 25, 1957.

W. H. DICKINSON is with the Esso Research and Engineering Company, Linden, N. J.

The author wishes to acknowledge the assistance of his associates in the preparation of this paper, particularly that of E. J. Winsor, R. J. Knopf, and C. L. Wilkinson.

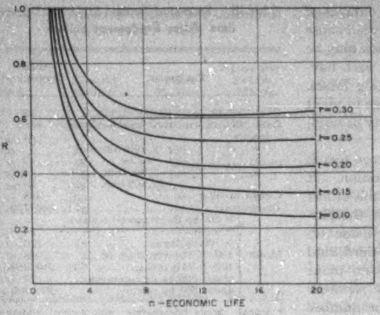


Fig. 1 (left). Values of R for $e=0.05$, $t=0.52$, $j_n=1/n$

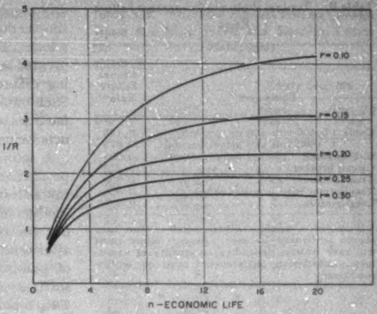


Fig. 2 (right). Values of 1/R for $e=0.05$, $t=0.52$, $j_n=1/n$, investment justified per dollar of gross earnings, E

In the use of equation 1, it is intended that the minimum acceptable return rate r be first determined. Knowing this and the other factors in equation 2, R can be calculated. After E is determined, I, the maximum justifiable investment, can be calculated.

Values of R have been calculated for various return rates and plotted against n, economic life, for a typical case when $e=0.05$, $t=0.52$, and $j_n=1/n$. These curves are shown in Fig. 1. Similarly, the inverse of this, values of $1/R$ are plotted against n, as shown in Fig. 2. This gives directly the investment justified per dollar of gross earnings, E, by letting E equal $1 in equation 1 or

$$I=1/R \tag{5}$$

A very important consideration in any economic analysis is what is an adequate return rate. This is often determined by company policy. In the absence of such a policy, a reasonable approach is to compare the risk of the proposed incremental investment with the company average return on net assets and with that of the plant or project, if known, weighing the relative (1) time pattern of earnings, (2) anticipated duration of earnings, and (3) certainty that the earnings will occur.

The per cent return on net assets of leading manufacturing corporations in 1955 varied between 5.5 and 29.1, the average being about 15. Generally speaking, return rates may be expected to be between 0.10 and 0.30 for most new construction in industrial plants.

Earnings that start immediately at a high level, continue for a long period, and are certain to occur without question, are typical of a lower-risk investment. On the other hand, earnings that are estimated to begin some time in the future, may not continue thereafter through the economic life of the investment, and which are less certain to occur, are of the higher-risk type. The lower-risk investment might reasonably have a somewhat lower return rate than the average, whereas, the higher-risk investment might have one a little higher to much higher than average.

To be specific, taking the average return rate of 0.15, the lower-risk investments might be in the 0.10–0.15 range, the medium-risk investment between 0.15–0.20, and the higher-risk investments in the 0.20-0.25 range. These values should be adjusted to the particular case under consideration.

Analysis of Power-System Reliability

The reliability of an industrial power system depends primarily on:

1. The quality of the supply and the supply system.
2. The quality of the system elements.
3. The circuit arrangement of the plant system.
4. The quality of the relay protective system.

Consideration should be given to each of

Table I. Values of a_n Where $a_n=r+\dfrac{r}{(1+r)^n-1}$

No. of Years, n	Rate of Return, r												
	0.10	0.12	0.14	0.15	0.16	0.18	0.20	0.22	0.24	0.25	0.26	0.28	0.30
1	1.100	1.120	1.140	1.150	1.160	1.180	1.200	1.220	1.240	1.250	1.260	1.280	1.300
2	0.576	0.592	0.6073	0.6151	0.6230	0.6387	0.6545	0.6705	0.6864	0.6944	0.7025	0.7186	0.7348
3	0.402	0.416	0.4307	0.4380	0.4453	0.4599	0.4747	0.4897	0.5047	0.5123	0.5199	0.5352	0.5506
4	0.3155	0.329	0.3432	0.3503	0.3574	0.3717	0.3863	0.4010	0.4159	0.4234	0.4310	0.4462	0.4616
5	0.2638	0.277	0.2913	0.2983	0.3054	0.3198	0.3344	0.3492	0.3643	0.3719	0.3795	0.3949	0.4106
6	0.2296	0.243	0.2572	0.2642	0.2714	0.2859	0.3007	0.3158	0.3311	0.3388	0.3466	0.3624	0.3784
7	0.2054	0.2191	0.2332	0.2404	0.2476	0.2624	0.2774	0.2928	0.3084	0.3163	0.3243	0.3405	0.3569
8	0.1874	0.2013	0.2156	0.2229	0.2302	0.2452	0.2606	0.2763	0.2923	0.3004	0.3086	0.3251	0.3419
9	0.1736	0.1877	0.2022	0.2096	0.2171	0.2324	0.2481	0.2641	0.2805	0.2888	0.2971	0.3141	0.3312
10	0.1628	0.1770	0.1917	0.1993	0.2069	0.2225	0.2385	0.2549	0.2716	0.2801	0.2886	0.3059	0.3235
11	0.1540	0.1684	0.1834	0.1911	0.1989	0.2148	0.2311	0.2478	0.2649	0.2735	0.2822	0.2998	0.3177
12	0.1468	0.1614	0.1767	0.1845	0.1924	0.2086	0.2253	0.2423	0.2597	0.2685	0.2773	0.2953	0.3135
13	0.1408	0.1557	0.1712	0.1791	0.1872	0.2037	0.2206	0.2379	0.2556	0.2645	0.2736	0.2918	0.3102
14	0.1358	0.1509	0.1666	0.1747	0.1829	0.1997	0.2169	0.2345	0.2524	0.2615	0.2707	0.2891	0.3078
15	0.1315	0.1468	0.1628	0.1710	0.1794	0.1964	0.2139	0.2317	0.2499	0.2591	0.2684	0.2871	0.3060
16	0.1278	0.1434	0.1596	0.1680	0.1764	0.1937	0.2114	0.2295	0.2479	0.2572	0.2666	0.2855	0.3046
17	0.1247	0.1405	0.1569	0.1654	0.1740	0.1915	0.2094	0.2278	0.2464	0.2558	0.2652	0.2843	0.3035
18	0.1219	0.1379	0.1546	0.1632	0.1719	0.1896	0.2078	0.2263	0.2451	0.2546	0.2641	0.2833	0.3027
19	0.1196	0.1358	0.1527	0.1613	0.1701	0.1881	0.2065	0.2252	0.2441	0.2537	0.2633	0.2826	0.3021
20	0.1175	0.1339	0.1510	0.1598	0.1687	0.1868	0.2054	0.2242	0.2433	0.2529	0.2626	0.2820	0.3016

Table II. Typical Probable Equipment Failure Rates

Equipment		Probable Annual Failure Rate*
Cable	2,400-volt and up, per mile	0.133
	600-volt and below, per mile	0.072
Cable joints		0.005
Circuit breakers		0.010
Transformers		0.003
Transformer with primary switch		0.004
Large motors		0.067

* Wherever available, actual service experience should preferably be used. Actual failure rates will vary widely, depending on quality of equipment, installation, maintenance, and on service conditions.

the foregoing points in determining the reliability requirements.

PROBABILITY

Based on service records of each type of equipment, the probability of failure of each may be found. Table II gives typical probable annual failure rates of equipment. These values may vary widely, depending on the quality of equipment and on installation conditions. Wherever available, actual service records should preferably be used.

In Table III the estimated forced outage times, or the electrical downtimes after equipment failure are given. Much has been written on the use of probability, but little is available in the industrial field.[4–6] By using probability methods, the relative reliability of different circuit or equipment arrangements may be determined.

A set of events is said to be independent if the occurrence of any one of them is not influenced by the occurrence of the others. The probability of failure of one kind of equipment is independent of that of another kind of equipment. Failure rates are generally expressed in terms of failures per year per unit of equipment. Whether or not a piece of equipment is in service has very little effect on the probability of failure. Although failures are not likely to occur when the equipment is out of service, incipient failure may develop during this period. It is assumed here that equipment failures are independent events, although there is some doubt on this point.[4] Using the principles given in Appendix II, the probability of forced outage, defined as the average number of unplanned out-of-service hours per year divided by 8,760, for any combination of elements or circuit arrangement may be found.

Generally speaking, a power outage may cause a loss in three ways. One is proportional to the number of outages, excluding, in many cases, the outages of shorter duration since these may not cause a loss. Those of longer duration may be divided according to duration, each having different values for resulting losses. Such outages may cause damage to plant facilities or may spoil the quality of products being manufactured.

A second kind of power outage loss is proportional to outage duration. It causes lost production. The value of lost production is real, except when the production requirements can be made up after service is restored. The third kind of power outage loss depends on plant startup time after service is restored. This loss is proportional to the number of outages and the plant start-up time.

Thus, the probable gross loss, E, which the plant sustains per year because of electrical failures is

$$E = fL_f + 8{,}760qL_p + fL_pT_s \qquad (6)$$

where

f = probable annual failure rate having outage time long enough to cause fixed loss
L_f = fixed loss per failure, f
L_p = production loss per hour of plant downtime
T_s = time required to start up plant
q = probability of being on forced outage

Take a simple example. If there are two elements of a system in series supplying a load, the failure of either will cause a forced outage. Assume that one is expected to fail once in 10 years and that 48 hours are required to effect repairs and to restore service. Assume that the other is expected to fail twice in 10 years and that 2 hours are required to effect repairs and to restore service. The combination will sustain three failures in 10 years and the service to the load will be interrupted for 52 hours in the 10-year period.

Table III. Estimated Typical Electrical Downtime After Equipment Failure

Equipment		Hours Downtime*
Cable failure, spare length on hand		28
Cable failure, run too long for spare length		41
Cable termination	2,400-volt and below	10
	6,900-volt	14
	26,400-volt	35+
Circuit breaker		4–8
Transformer		360–720
Motor	0–100 hp (horsepower) replacement	24
	0–100 hp repair	170
	100–500 hp replacement	48
	100–500 hp repair	240–360
	above 500 hp replacement	72
	above 500 hp repair	400–600

* Actual hours downtime will vary widely. Wherever available actual service records should be used.

If we let

$f_a = 1/10 = 0.10$, probable annual failure rate of first element
$f_b = 2/10 = 0.20$, probable annual failure rate of second element

then,

$$f_{\text{total}} = f_a + f_b \qquad (7)$$

$$= 0.10 + 0.20 = 0.30 \qquad (8)$$

Similarly, let

S_a, S_b = shutdown or forced outage hours per failure of first and second elements respectively
q_a, q_b = probable annual shutdown or forced outage hours per year divided by 8,760, that is, the probability of being on forced outage

Table IV. Assumed System Probable Failure Rates (1960–1969)

Item	System Elements: Description	Quantity	f, Per Unit	Total f
	Case 1—Simple Radial System. See Fig. 3			
a_1	Supply	1	3.000	3.000
a_2	Supply	1	0.100	0.100
Total a				3.100
b_1	Switch	1	0.001	0.001
b_2	Feeder circuit breaker	1	0.010	0.010
b_3	Feeder cable	1/2 mile	0.133	0.067
Total b				0.078
c_1	Transformer and switch	1	0.004	0.004
c_2	Secondary circuit breaker	1	0.010	0.010
Total c				0.014
Total				0.192 or 3.192
	Case 4—Primary-Selective System. See Fig. 6			
b_1	Power-supply switch	1	0.001	0.001
b_2	Switch transfer	1	0.244	0.178
c	Total			0.014
Total				0.259
	Case 6—Secondary-Selective System. See Fig. 8.			
c_3	Tie circuit breaker	1	0.010	0.010
c_4	Breaker transfer	1	0.258	0.192
Total				0.202

Table V. Assumed Outages and Losses per Failure (1960–1969)

Item	Element S, Hour	Plant T_s, Hour	Loss Per Failure, M\$, L_f
Case 1—Simple Radial System. See Fig. 3			
a_1	0.5	0	0
a_2	24	20	5
b_1	48	20	5
b_2	8	20	5
b_3	48	20	5
c_1	500	20	5
c_2	8	20	5
Case 4—Primary–Selective System. See Fig. 6			
b_1	48	20	5
b_2	1	0	0
c_1	500	20	5
c_2	8	20	5
Case 6—Secondary–Selective System. See Fig. 8			
c_3	8	20	5
c_4	0.000+	0	0

Now,

$$q_a = f_aS_a = 0.10 \times 48/8,760 \tag{9}$$

$$q_b = f_bS_b = 0.20 \times 2/8,760 \tag{10}$$

Hence,

$$q_{\text{total}} = (4.8+0.4)/8,760 = 5.2/8,760 \tag{11}$$

$$= 0.000594 \tag{12}$$

If alternate supplies or circuits are available to a load, the failure rate will be greater because of the added complexity, but the probability of being on forced outage will usually be reduced because of the reduction in time required to restore service. In Appendix II, it is seen that squared q terms, or products of q terms, are generally negligible. In what follows, it is assumed that such terms are negligible.

Typical Case Study

Using the principles just described, the general equations of power system reliability for six simplified but typical systems are developed in Appendix III. For each of the six, reliability is expressed in terms of (1) the probable number of failures per year, and (2) the fraction of the year that the system is likely to be on forced outage, that is, the probability of forced outage.

Example

To illustrate the application of the general equations developed in Appendix III, a simple example is worked out. In Table IV, typical failure rates are assigned to the elements of systems shown in Figs. 3, 6 and 8, and the probable total system failure rates are calculated.

Table VI. Assumed Plant Hourly Production Loss

Years	Hourly Loss, M\$, L_p
1960–1964	1.0
1965–1969	1.3

Table V lists (1) assumed outages resulting from equipment failures, both to repair the failure and restore electrical service, and to restart the plant; and (2) the fixed loss per failure. Table VI gives values assumed for plant hourly production loss during forced power system outages plus plant startup time after a failure. It was assumed that plant production losses increase after the first 5 years. This may arise because of an expected increase in unit value of product or in production.

Table VII gives probable annual power system forced outage time because of failures of equipment, and the yearly plant fixed and start-up losses. Table VIII summarizes the total annual losses that result from system outages. These losses are made up of three elements: (1) plant production losses, (2) fixed plant losses, and (3) plant startup losses.

The equivalent uniform annual value, or levelized value of the gross earnings (or losses in these cases), E, may be found by the equation

$$E = a_n \sum_{n=1}^{n=n} E_n/s_n \tag{13}$$

In this equation it is assumed that the investment is made at the beginning of 1959 one year before the plant begins production. The system gross losses, as given in Table VIII, may be levelized as indicated in equation 13, where n, the economic life of the investment is 11 years, the plant being in production for 10 years. However, because the losses are relatively uniform, it is simpler to levelize them as shown in Appendix IV.

Table VII. Probable Annual Forced Outage Times and Plant Fixed and Start-Up Losses (1960–1969)

Item	fS (8,760 q)	Yearly Loss M\$ Fixed fL_f	Yearly Loss M\$ Start-Up fT_sL_p
Case 1—Simple Radial System. See Fig. 3			
a_1	1.50	0	0
a_2	2.40	0.50	2.0
b_1	0.05	0.005	0.02
b_2	0.08	0.05	0.20
b_3	3.36	0.35	1.40
c_1	2.00	0.02	0.08
c_2	0.08	0.05	0.20
Total	9.47	0.975	3.90
Case 4—Primary–Selective System. See Fig. 6			
b_1	0.05	0.005	0.02
b_2	0.18	0	0
c	2.08	0.070	0.28
Total	2.31	0.075	0.30
Case 6—Secondary–Selective System. See Fig. 8			
c_3	0.08	0.4	0.20
c_4	0	0	0
Total	0.08	0.4	0.2

Next, in Appendix IV, the different gross earnings of the three power systems are calculated. The investment charge factor, R, of equation 2 is then calculated. It was found to be 0.423 for the values assumed. Lastly, in Appendix IV, the amounts that can be justified to improve power-system reliability in this typical example are calculated. In an actual system study, the number of system elements will usually increase and actual values may be much different from those assumed in the example.

Conclusions

Equation 1 may be used to determine the incremental investment that is justified when it produces gross earnings E. Through the use of probability principles, the probability of power system failures and forced outages may be calculated. General equations for power systems reliability are given in Appendix III. The value of plant losses resulting from forced outages may be calculated using the principles of these general equations, assigning suitable values of equipment failure rates, forced outage time per failure, fixed plant loss per failure, and production loss per hour of forced outage and per hour of plant start-up time.

With this information, equation 1 may be used to determine how much increased

Table VIII. Summary of Annual Plant Losses from Forced Outages

System	Year	Yearly Shutdown Losses, M\$ Electrical Failure, 8,760 q = fSL_p	Fixed, fL_f	Start-Up fT_sL_p	Total (−E)
Simple radial, Case 1	1960–1964	9.47	0.975	3.90	14.35
	1965–1969	12.30	0.975	5.07	18.35
Primary selective, Case 4	1960–1964	2.31	0.075	0.30	2.69
	1965–1969	3.00	0.075	0.39	3.47
Secondary selective, Case 6	1960–1964	0.08	0.400	0.20	0.68
	1965–1969	0.10	0.400	0.26	0.76

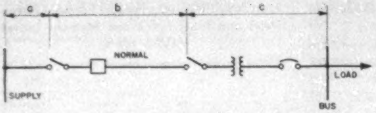


Fig. 3. Case 1—Simple radial system

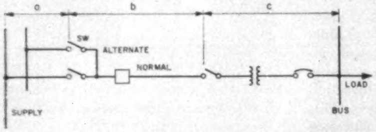


Fig. 4. Case 2—Radial system with selective independent supplies

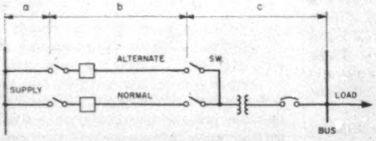


Fig. 5. Case 3—Primary selective system with single supply source

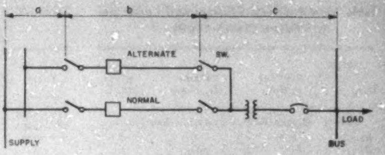


Fig. 6. Case 4—Primary-selective system with independent supplies

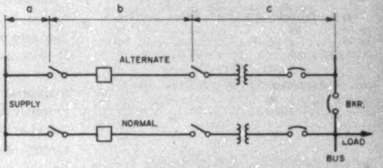


Fig. 7. Case 5—Secondary-selective system with single supply source

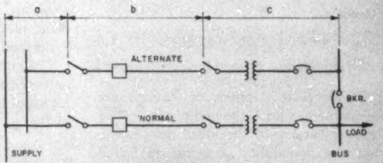


Fig. 8. Case 6—Secondary-selective system with independent supplies

investment is justified to improve power-system reliability. The technique described in this paper enables the question to be answered: Is the extra investment required for a secondary-selective system over a radial system justified, or is the use of a higher quality system element justified?

Appendix I. Derivation of Equation for Justifiable Investment

Annual net income is found by deducting from the gross revenues all operating expenses, depreciation, maintenance, state taxes, insurance, and income taxes. Thus, net income is

$$N=(G-X-eI-j_nI)(1-t) \tag{14}$$

Since cash flow is defined as net income plus depreciation

$$C=N+D \tag{15}$$

$$=(G-X-eI-j_nI)(1-t)+j_nI \tag{16}$$

But C is an annual series of sums which will give a return rate on the investment I plus an amount compounded at rate r for n years, the economic life of the project, sufficient to recover the investment. Hence:

$$a_n=C/I \tag{17}$$

the capital recovery factor. Substituting $C=a_nI$ into equation 16, and solving for I, there results then the following equation:

$$I=E/R \tag{18}$$

where:

$$R=\frac{a_n-j_nt}{1-t}+e \tag{19}$$

$$E=G-X \tag{20}$$

Appendix II. Probability Principles

For a set of, say, two elements in series, the failure of either one of which will take the series out of service, the probability that both will remain in service, or will not fail, is $p=p_1p_2$; where $p_1=(1-q_1)$, $p_2=(1-q_2)$, and q_1q_2 is the probability of forced outage, that is, $q_n=f_nS_n/8{,}760$. In this, f is the probable annual failure rate, S_n is the average number of hours the element, circuit, or plant is out of service because of the failure. Thus, the probability of a forced outage of either of the elements is

$$q=1-p_1p_2,\text{ or } q=1-(1-q_1)(1-q_2)$$

$$=(q_1+q_2)+q_1q_2$$

When q_1 and q_2 are less than 0.01, the error in dropping the term q_1q_2 is less than one per cent. Since this applies to most probabilities of circuit or equipment element forced outage, it is generally sufficiently accurate to combine probabilities of forced outage by adding the forced outage rates, and to neglect the terms involving products.

For two elements that are in parallel, whose forced outage probability are q_1 and q_2 respectively, the probability that both will be out of service simultaneously is q_1q_2. The probability that both will be in service is $p=(1-q_1)(1-q_2)$ or $p=1-(q_1+q_2)+q_1q_2$. The probabilities of all possible operating conditions of two elements is given by (p_1+q_1) $(p_2+q_2)=p_1p_2+p_1q_2+p_2q_1+q_1q_2=1$. In this, the probability of being in service with both elements is p_1p_2. The probability of being out of service with both elements is q_1q_2. The probability of being in service with only one element is $p_1q_2+p_2q_1$.

The events, element *1* in service and element *2* out of service, and the event, element *1* out of service with element *2* in service, are known as mutually exclusive events. When events are mutually exclusive, they cannot occur simultaneously and the occurrence of one prevents the other. The probability of the occurrence of either one of the two mutually exclusive events is equal to the sum of the probabilities of the single events.

Appendix III. General Equations for Power System Reliability

In the following refer to Figs. 3–8. These cover six simplified typical power systems. In the following equations, the postsubscripts *a*, *b*, *c*, refer to system elements, and *r*, *p*, *s*, to type of system. Presubscripts refer to the case under study. Symbols used in reference to alternate supplies or circuit elements are primed. Symbols referring to normal supplies or circuit elements are unprimed.

Referring to Fig. 3, a simple radial system (Case 1) is considered. The probability of annual system failures resulting in forced outages is

$$ {}_1f_r=f_a+f_b+f_c \qquad (21)$$

The probability of the system being on a forced outage is

$$ {}_1q_r=q_a+q_b+q_c \qquad (22)$$

Similarly, for Case 2, a radial system with selective independent supplies, as shown in Fig. 4, the probability of failures and of forced outages is respectively

$$ {}_2f_r={}_1f_r+f_{sw}'+f_a'q_a \qquad (23)$$

$$ {}_2q_r=q_a^2+q_b+q_c+f_{sw}'(S_{sw}'/8,760)+f_a(S_{\text{transfer}}/8,760) \qquad (24)$$

$$ \approx q_b+q_c+q_{sw}'+q_{\text{transfer}} \qquad (25)$$

For Case 3, a primary-selective system with a single supply source, as shown in Fig. 5, the corresponding equations are

$$ {}_3f_p={}_1f_r+f_{sw}'+(f_b')(q_a+q_b) \qquad (26)$$

$$ {}_3q_p=q_a+q_b^2+q_c+q_{sw}'+f_b(S_{\text{transfer}}/8,760) \qquad (27)$$

$$ \approx q_a+q_c+q_{sw}'+q_{\text{transfer}} \qquad (28)$$

For Case 4, a primary-selective system with independent supplies, shown in Fig. 6:

$$ {}_4f_p={}_1f_r+f_{sw}'+(f_a'+f_b')(q_a+q_b) \qquad (29)$$

$$ {}_4q_p=(q_a+q_b)^2+q_c+q_{sw}'+q_{\text{transfer}} \qquad (30)$$

$$ \approx q_c+q_{sw}'+q_{\text{transfer}} \qquad (31)$$

In Case 5, a secondary-selective system with a single supply, shown in Fig. 7:

$$ {}_5f_s={}_1f_r+f_{bkr}+(f_b'+f_c')({}_1q_r) \qquad (32)$$

$$ {}_5q_s=q_a+(q_b+q_c)^2+q_{bkr}+(f_b+f_c)(S_{bkr}/8,760) \qquad (33)$$

$$ \approx q_a+q_{bkr}+q_{\text{transfer}} \qquad (34)$$

Finally, in Case 6, a secondary-selective system with independent supplies shown in Fig. 8, the corresponding equations are

$$ {}_6f_s={}_1f_r+f_{bkr}+f_r'({}_1q_r) \qquad (35)$$

$$ {}_6q_s={}_1q_r^2+q_{bkr}+{}_1f_r(S_{\text{transfer}}/8,760) \qquad (36)$$

$$ \approx q_{bkr}+q_{\text{transfer}} \qquad (37)$$

Appendix IV. Computation of Levelized Annual Losses, Differential System Earnings, Investment Charge Factor and Justifiable Investment

Levelized Annual Losses

For Case 1—Radial system. See Fig. 3 and Table VIII.

$$ -E_1-=14.35+\frac{(18.35-14.35)(a_{11})}{a_1s_4}-a_{11}\frac{14.35}{s_1} \qquad (38)$$

$$ =14.35+\left(\frac{4.00}{0.334(2.986)}-\frac{14.35}{1.20}\right)\times(0.2311) \qquad (39)$$

$$ =\text{M\$ }12.51 \qquad (40)$$

For Case 4—Primary-selective system. See Fig. 6 and Table VIII.

$$ -E_4=2.69+\left(\frac{0.78}{0.334(2.986)}-\frac{2.69}{1.20}\right)(0.2311) \qquad (41)$$

$$ =\text{M\$ }2.35 \qquad (42)$$

For Case 6—Secondary-selective system. See Figure 8 and Table VIII.

$$ -E_6=0.68+\left(\frac{0.08}{0.334(2.986)}-\frac{0.68}{1.20}\right)(0.2311) \qquad (43)$$

$$ =\text{M\$ }0.567 \qquad (44)$$

Differential System Earnings (Negative Losses)

For Case 4 over Case 1

$$ E_{4-1}=E_4-E_1 \qquad (45)$$

$$ =\text{M\$ }(-2.35+12.51) \qquad (46)$$

$$ =\text{M\$ }10.16 \qquad (47)$$

Similarly for Case 6 over Case 1

$$ E_{6-1}=\text{M\$ }(-0.567+12.51) \qquad (48)$$

$$ =\text{M\$ }11.94 \qquad (49)$$

And for Case 6 over Case 4

$$ E_{6-4}=\text{M\$ }(-0.567+2.35) \qquad (50)$$

$$ =\text{M\$ }1.78 \qquad (51)$$

Calculation of the Investment Charge Factor

The value of *R* in equation 2 is

$$ R=\frac{r}{1-t}+e+\frac{d_n-j_nt}{1-t} \qquad (52)$$

The following values are assumed as typical for the plant investment,

$$ \begin{aligned} r&=0.20\\ t&=0.52\\ e&=0.050\\ n&=11\\ j_n&=\frac{1}{n-1}=0.1 \end{aligned} \qquad (53)$$

Substituting these values, there results

$$ d_n=\frac{0.20}{1.20^{11}-1} \qquad (54)$$

$$ =0.031 \qquad (55)$$

$$ R=\frac{0.20}{1-0.52}+0.50+\frac{0.031-0.1\ (0.52)}{1-0.52} \qquad (56)$$

$$ =0.417+0.050-0.044 \qquad (57)$$

$$ =0.423 \qquad (58)$$

Computation of Justifiable Investment

The incremental investment justified to obtain increased power system reliability is, from equation 1, for Case 4 over Case 1,

$$ I_{4-1}=E_{4-1}/R \qquad (59)$$

$$ =\text{M\$ }10.16/0.423 \qquad (60)$$

$$ =\text{M\$ }24.0 \qquad (61)$$

For Case 6 over Case 1,

$$ I_{6-1}=\text{M\$ }11.94/0.423 \qquad (62)$$

$$ =\text{M\$ }28.2 \qquad (63)$$

For Case 6 over Case 4,

$$ I_{6-4}=\text{M\$ }1.78/0.423 \qquad (64)$$

$$ =\text{M\$ }4.20 \qquad (65)$$

References

1. Principles of Engineering Economy (book). E. L. Grant. The Ronald Press, New York, N. Y., 1930, 1938, and 1950.

2. Measuring the Productivity of Capital, Joel Dean. *Harvard Business Review*, Cambridge, Mass., Jan.-Feb. 1954, pp. 120–30.

3. Interest Rate of Return for Capital Expenditure Evaluation, James B. Weaver, Robert J. Reilly. *Chemical Engineering Progress*, New York, N. Y., vol. 52, no. 10, Oct. 1956, pp. 405–12.

4. The Use of Probability in the Design and Operation of Secondary Network Systems, N. M. Neagle, D. R. Nelson. *AIEE Transactions*, vol. 74, pt. III, Aug. 1955, pp. 567–73.

5. Higher Mathematics for Engineers and Physicists (book), I. S. and E. S. Sokolnikoff. McGraw-Hill Book Company, Inc., New York, N. Y., 1934.

6. The Probability of Coincident Primary-Feeder Outages in Secondary Networks, David N. Reps. *AIEE Transactions*, vol. 73, pt. III-B, Dec. 1954, pp. 1467–78.

7. Coincident-Outage Probability in Secondary-Network Vaults, David N. Reps. *Ibid.*, vol. 74, pt. III, Aug. 1955, pp. 580–86.

8. Probability Calculations for System Generation Reserves, Carl Kist, G. J. Thomas. *Electrical Engineering*, vol. 75, Sept. 1956, pp. 806–11.

◆

Discussion

B. E. Crawford (Esso Standard Oil Company, Baton Rouge, La.): Mr. Dickinson's paper is an excellent comprehensive approach to the problem facing the electrical engineer today in deciding to what extent money should be spent to provide additional increments of reliability to a particular load. At Esso's Baton Rouge Refinery similar economic evaluations have been performed for some time. This paper provides briefly most of the considerations necessary to fully evaluate not only the factual data but the proper assumptions and estimates to arrive at a practical solution as to the worth of the electric system to the particular industry. With the high cost of downtime in many industries, information such as this will show many engineers that a more modern, reliable electric system is often justified.

Although this paper deals primarily with the reliability of the electric system, it gives no consideration to maintenance. The quality of the system is very dependent on maintenance, and the degree of maintenance is usually closely associated with the system design. With large systems in continuous operation, maintenance considerations may be a major concern in system design. Therefore, in determining system design through economic evaluation, consideration must also be given to how maintenance will be performed. For instance, one of the economic benefits derived from a secondary selective-type substation is the ease of performing certain types of maintenance with no loss in production.

The author points out the idea of independent events as associated with probability. The experience at the Baton Rouge Refinery indicates there is much doubt that the type of equipment failures under consideration here are always independent events. However, such dependent failures are usually due to unknown or unpredictable factors, and it is not practicable to consider such failures. We normally design on the basis that we will not have two emergency conditions to occur during the same period of time on the same system.

One of the most difficult items to derive for use in economic evaluation is probable equipment failures. As stated in the paper, actual records of such failures are the best source of information. However, we have found that constant changes to new and better equipment in new construction does not provide long time records suitable for accurately determining this information. We have also felt that this one factor has caused us to put less emphasis on actual failure rate. It would be interesting to know the source of the failure rates given in Table II.

We feel that the method outlined here is the most acceptable method we know of in helping to evaluate investments. In addition to this method, however, we feel past experiences and good judgment definitely should be used in the final analysis as to how much additional money should be expended for improved reliability.

D. S. Brereton (General Electric Company, Schenectady, N. Y.): This paper contains one of the most complete presentations of the economic evaluation of the modern industrial system. One of its principal attributes is that it draws attention to the fact that such an evaluation is possible. The use of the principles of Mr. E. L. Grant have been effectively modified for the power system field.

One of the important concepts that relate to the final analysis is the material given in Table II. It is the user, and not the manufacturer, that has the most complete information on the performance of equipment. It appears that the numbers included in this Table are relatively in the right proportion with the exception of the material dealing with cable joints and circuit breakers. I believe it possible that these two numbers are reversed. It would further be helpful to note, under the subject of transformers, whether liquid-filled or dry-type units are concerned. Although the transformer has the lowest failure rate, if the number given applies to a liquid-filled transformer, the dry-type transformer would probably be in the neighborhood of 0.004.

Mr. Dickinson has done a very capable job of effectively grouping together material on this subject. It is hoped that the AIEE Industrial Power Systems Committee will further document and expand the material given in this paper.

H. B. Backenstoss (Jackson & Moreland, Inc., Boston, Mass.): The author has undertaken a commendable and difficult task in isolating and defining in dollar terms the specific item of reliability in an industrial electric system. He has made an important contribution on an important subject.

Although the application of probability to electric system design and the evaluation of outage losses are the aim of this paper, the section dealing with the technique for economic evaluation has the prominence of first appearance. The development of the formulas for this technique raises some questions which bear examination even at the expense of neglecting other aspects of this worthwhile paper. The following discussion is confined entirely to the evaluation technique.

The discounted-cash-flow method used by the author and described in several of his references is a mathematically rigorous method for determining the interest rate of return. Basically it is a statement of the identity underlying all money lending calculations, namely, that the present worth of all future interest payments plus repayments of principle must exactly equal the initial value of the loan. In the discounted-cash-flow method of analyzing an investment, the earnings on the investment (interest) and annual depreciation accruals (repayments of principal) are lumped together as total cash flow and not individually identified. The interest rate for the present worth arithmetic which satisfies the identity of these payments with the original investment is the rate of return earned by that investment.

A project's cash flow has the advantage of being a definite measurable amount. It very usefully by-passes the problem of depreciation which has caused endless confusion as to which should be used, i.e., tax-allowable, straight line, or sinking fund. It can be correctly determined by two different computations, either by steps typical of those used by a company in computing its Federal Income tax obligation, or by steps typical of those it uses in preparing its income statement for stockholders. Due to the possibility open to a company to use accelerated depreciation in determining its taxable income and normal depreciation in determining its reported income, the two methods are not identical, although they both show exactly the same cash flow, i.e., net income plus depreciation.

The author has used the "tax form" approach in developing expressions for *R* in the text and in Appendix I. This conceals the tax credit which the difference between tax allowable depreciation and economic depreciation (sinking fund) provides. The importance of this difference is increased where a company uses SYD (sum-of-years digits) or other similar accelerated methods for tax purposes.

The expressions for *R* may readily be transposed to a form which reveals the tax credit, as follows:

$$R=\frac{a_n-j_nt}{(1-t)}+e \tag{66}$$

$$=\frac{a_nt}{(1-t)}+\frac{a_n(1-t)}{(1-t)}-\frac{j_nt}{(1-t)}+e \tag{67}$$

$$=a_n+\frac{t}{(1-t)}(a_n-j_n)+e \tag{68}$$

$$=r+d_n+e+\frac{t}{(1-t)}[r-(j_n-d_n)] \tag{69}$$

This expression separates return r from the other components of total carrying charge on investment and also identifies income tax as a separate item in terms of the tax rate t, the return r and a credit related to the difference between tax-allowable depreciation and economic depreciation. Expense e and depreciation d_n are the same as in the author's equation 2. These expressions would all be somewhat different for a company which had substantial debt financing.

It may appear on the surface that, if this formula gives the same result for *R* as the author's, the distinction between the two is of no moment. In fact, however, there are subtle overtones of some importance.

It is desirable that an economic evaluation, which aims at measuring the earning power of a proposed new investment, adopt the methods which the company itself uses in measuring its own earning power. To use other methods in evaluating a project which, after acceptance, becomes an increment addition to the company's property is apt to be confusing. This is particularly true in the case of management, which is almost surely predisposed to think in terms of the company's own methods when reviewing proposed capital expenditures.

The methods used by a company to measure its earning power are the methods used in preparing an annual income statement, for those methods, although not standard for all companies and subject to discretionary variation within a given company, are the ones used to determine net income. And net income as it appears on a company's income statement is generally regarded as the measure of a company's profitability. Net income as it appears on a tax statement is commonly a much lesser amount. It fails to include the depreciation tax credit exemplified in the last term of equation 4 above. The author's equation 14 defines net income as it appears on a tax statement. In reality, it defines "taxable income less taxes," a quite different figure from net income.

By taking the "income statement" approach in evaluating proposed capital investments one avoids falling into the error of thinking that income taxes should be 52% (the current rate) of income before taxes. The 1956 annual report of Standard

Oil Company of New Jersey furnishes an example:

Net income before taxes	$1,238,500,000	100.0%
Provision for income taxes	430,000,000	34.7%
Net income	$ 808,500,000	65.3%

From equation 14 one might expect income taxes to be 52% of net income before taxes, rather than only 34.7%. Accelerated tax-allowable depreciation is almost surely the primary cause of the great reduction in the tax that is evident here.

One also avoids another pitfall, a second facet of the same problem. It is the pitfall of interpreting the ratio of taxable income less taxes divided by investment as the rate of return r. Referring again to Standard Oil Company's income statement:

Provision for income taxes	$430,000,000	52%
Corresponding taxable income (computed as $430,000,000/0.52)	827,000,000	100%
"Net income" (Actually taxable income less taxes)	$397,000,000	48%
Actual net income	$808,500,000	

The tax credit provided by accelerated depreciation adds directly to net income and is a significant amount. It would be quite incorrect to use the much smaller item of taxable income less taxes as a gauge of profitability.

By separating cash flow a_n into its components, return r and depreciation d_n, in equation 69, one can also see that a plan of low cash flow, but low depreciation due to long life, may provide a higher return and hence be more attractive than a plan of higher cash flow having a high depreciation. This clarifies the author's statement about cash flow. It is not the flow itself but the interest rate of return it provides which determines the economic choice of plans, and that is what he means.

In this connection, it is well to bear in mind that the rates of return between 5.5 and 29.1% reported by the author are not representative values to use for new investments. They are too high, having been inflated by past price rise. They are the earnings rates at current price levels on investments committed at lower price levels. The amount of increase in earnings rate caused by this effect may be difficult to determine. It would appear, however, that 15%, the author's suggested average rate of return, is appreciably higher than the rate at which past investments were made and were believed attractive.

Another factor inherent in the development of the formulas in this paper will probably be evident to most of those who make economic comparisons. By treating only the differential investment and differential annual costs, the formulas presuppose that the carrying charge rate for the increment is the same as the carrying charge rate for the base investments. Although this is probably the case in most electric design alternatives it need not invariably be so. One plan may have a longer life or different degree of risk and its total carrying charge rate will differ from that of a second plan. In this event the two plans cannot accurately be compared by the formulas developed here. Where the carrying charge rate varies for any reason it becomes necessary to calculate annual charges for each plan on either a year-by-year or a levelized basis and obtain differences in cost by direct subtraction.

The author is to be congratulated for his interest and his efforts in placing the economic evaluation of reliability in an electric system on a sound mathematical basis, instead of a basis of personal preferences and guesswork which have often heretofore been the principal guiding influences.

W. H. Dickinson: The author wishes to express his appreciation to Mr. Crawford, Mr. Brereton, and Mr. Backenstoss for their discussions. The point raised by Mr. Crawford regarding the importance of maintenance is a good one. Contrary to his impression, the equation $I=E/R$, developed in the paper, does take maintenance into account. In addition to normal maintenance, even losses resulting from plant downtime which are required to make additions to the electric system, may be accounted for. The value of E, the gross earnings, equals gross revenues less variable expenses X. Thus X covers variable maintenance expenses as well as other variable expenses. The investment charge factor R includes the expense rate e, which covers fixed expenses, including fixed maintenance. Of course, if the values of variable maintenance are not uniform during the life of the project, they must be levelized as described in the paper. Essentially, this consists of finding the uniform annual amounts over the life of the project having a present worth equal to the present worths of the variable maintenance costs.

Mr. Crawford also questioned whether the probability of forced outages resulting from system component failure should be regarded as independent events. After wrestling with this problem for some time, the author concluded that it was more appropriate to treat forced outages as independent events. This is the method used recently[1] in calculating forced outage rates on unit-type installation of boiler and turbine generators. In a discussion of their paper (see reference 4 of the paper) Mr. Neagle and Mr. Nelson stated that they now regarded these events as mutually exclusive, whereas in a previous paper they had treated them being independent events. The author believes that forced outages are not truly mutually exclusive in that the occurrence of one forced outage does not prevent the occurrence of a simultaneous failure.

From a practical point of view, the forced outages rates of electric equipment in industrial power systems are usually low. The forced outage rates of most simple radial systems will usually be less than one per cent. Hence, it will make no significant difference in final results whether forced outage rates or failure rates are treated mathematically as independent events, or as mutually exclusive events. For example, let forced outage rates of one per cent be assigned to two items connected in series. If the probability of failures are regarded as mutually exclusive events, the combined forced outage rate will be found accurately by addition of the forced outage rates, that is, $0.01+0.01=0.02$, or 2%. If, instead, the events are regarded as independent, the probability of being in service is found by multiplying in service probabilities, that is, $(1-0.01)(1-0.01)=1-2(0.01)+(0.01)(0.01)=1-0.02+0.001$. The probability of being on forced outage is found by subtracting this term from one, giving the result $0.02-0.0001$. Dividing the difference, 0.0001 by 0.02 shows that there is only 0.5% difference in the two results. This is well within the accuracy of the forced outage rates themselves.

As Mr. Crawford points out, economic evaluation is not the only consideration in determining power system reliability requirements. There are always intangibles which cannot be evaluated. But past experience should not be one of them, because it should be possible to measure past experience, that is, system failures and their effects on plant earnings, in dollars and cents. If this is done, preconceived notions of what can be justified will be avoided.

Mr. Brereton and Mr. Crawford have both commented on Table II, "Typical Probable Equipment Failure Rates." The sources of these data are (1) published reports of the Edison Electric Institute and given in reference 4 of the paper, (2) AEIC (Association of Edison Illuminating Companies) Report of Committee on Power Distribution 1954–1955, covering cable operation for the period 1943–1953, and (3) scattered and incomplete information from various other sources. None of these sources included information on circuit breaker failure rates. These were estimated. However, as a check, the Edison Electric Institute gives a failure rate of network protectors to close as 0.0292. Failure of a network protector to close is considered more likely than the failure of a circuit breaker. Hence, a circuit breaker failure rate of 0.01 was selected. As to cable joints, failure rates of 0.005 and 0.00159 were reported by the EEI and the AEIC reports respectively. Similarly, failure rates for cable per mile per year are given as 0.068 and 0.072 by the EEI and AEIC reports respectively. In each case, the reported figures were arbitrarily increased to those quoted in Table II, to allow for industrial conditions, which were thought to be more severe than the utility experience on which the EEI and AEIC reports were based. Transformer failure rates were based on the EEI report in which the figure 0.00274 was quoted. These are undoubtedly liquid filled transformers. In Table II this figure was rounded out to 0.003, for a transformer only and to 0.004 for a transformer with a switch. The author has no information on dry-type transformers, but Mr. Brereton's suggested failure rate of 0.004 appears reasonable.

It is unfortunate that there is so little information available from industrial sources. With some digging no doubt some information could be found in the records of maintenance departments. Over-all forced outages of plants should also be available from the record of plant operating personnel, which could be used to check results. The author concurs wholeheartedly with Mr. Brereton's suggestion that the AIEE Industrial Power Systems Committee further document and expand the material given on the paper. A practical way of doing this might be the preparation of a questionnaire similar to the form prepared on forced outage rates on steam turbines and boilers.[1]

Mr. Backenstoss has confined his discussion to the economic aspects of the paper. He agrees that the discounted-cash-flow method is a mathematically rigorous method for determining the interest rate of return.

Regarding depreciation, far from bypassing, as Mr. Backenstoss indicates, the problem of depreciation, the paper attempts to meet this problem head-on. It provides for the use of sinking-fund depreciation for the economic analysis after computing income taxes, using for tax purposes the type of depreciation actually used by the company in question, and permitted by the income tax authorities. It is sometimes the case that a company uses still a third type of depreciation for financial accounting purposes. For example, it may use accelerated depreciation, such as sum-of-years' digits or declining balance, for income tax purposes, and normal or straight line depreciation for financial accounting purposes. In such cases, a tax equalization reserve or reserve for deferred income taxes is often set up. This third type of depreciation is of no consequence in an economic analysis. The term

$$N=(G-X-eI-j_nI)(1-t) \tag{70}$$

as given in equation 14 of the paper precisely defines the meaning of N. If the term, j_nI, the depreciation used in computing income taxes is also used for financial accounting purposes, and if no capital for the incremental investment I was derived from borrowing, then, and only then, may the term N be properly called "net income after taxes." In the paper this was assumed to be the case for the incremental investment and hence N was called net income. However, this is a loose use of the term, which, as used by accountants, means the amount earned or available for the stockholders after allowing for income taxes, interest payments to lenders, and book depreciation. Book depreciation may differ from that used in computing income taxes. There is no accepted term to describe N. Perhaps "taxable income less income tax (no borrowed funds)" would be a better way to refer to N. Similarly, equation 15 in the paper precisely defines cash flow. It is

$$C=N+D \tag{71}$$

where N is as defined above and D is the j_nI, the income tax depreciation. Neither the question of semantics in referring to N, nor the treatment of the incremental investment on the books of account should be allowed to confuse the issue when making an economic analysis. For this, find the annual cash flows C over the life of project, and from these determine r the rate of return on the investment, as described in the paper.

Mr. Backenstoss states that cash flow, as determined by equation 2, will be the same as net income after taxes plus book depreciation where book depreciation differs from income tax depreciation. This is not correct. For example, assume that the installed cost of a device is $300, has a 3-year life, and has gross earnings (before depreciation or income tax) of $200. The first year the straight line depreciation is $100 and the declining balance depreciation is $200. If the income tax rate is 50%, the cash flow in the former case is

$$C=(200-100)(0.50)+100 \tag{72}$$

$$=\$150 \tag{73}$$

and in the latter case is

$$C'=(200-200)(0.50)+200 \tag{74}$$

$$=\$200 \tag{75}$$

Mr. Backenstoss has developed in his equation 69 another form for the investment charge factor than those shown in the paper where the two different forms were used:

$$R=\frac{a_n-j_nt}{1-t}+e \tag{76}$$

$$R=\frac{r}{1-t}+e+\frac{d_n-j_nt}{1-t} \tag{77}$$

The form developed by Mr. Backenstoss is as follows:

$$R=r+d_n+e+\frac{t}{1-t}\,[r-(j_n-d_n)] \tag{78}$$

All three of these are algebraic identities. It would appear to be a matter of personal preference which one is used. The author prefers the first form because it is the most concise and in the best form for ease in computation. It is also meaningful when the significance of the terms is understood.

All three of these forms give exactly the same results and therefore give the same measure of the earning power or profitability. The methods described in the paper are intended for use in determining the incremental investment that can be justified. They are not intended for measuring the profitability of over-all company operations. In any event, it is not net income after income taxes but the annual cash flows, as defined in equation 2, or equation 16 in the paper, over the life of the project which should be used in an economic analysis to determine the rate of return. This is at the very heart of the discounted-cash-flow method which Mr. Backenstoss agrees is a mathematically rigorous method.

The paper assumes that investment in the project is not from borrowed funds, that is, that no interest payments are involved. This simplifies the economic analysis and avoids the necessity of determining the percentage of borrowed capital and the interest rates being paid thereon. Because of the tax credit on interest payments, the rate of return is higher when a part of the funds are borrowed. A correction[2] can be made to the rate of return, if a fraction of the investment is actually derived from borrowing. In an industrial company, it is usually the case that the fraction of the investment derived from borrowing is lower, and the rate of return on investment is higher, than that which prevails in public-regulated utilities. As a consequence, the magnitude of the correction for borrowing is usually smaller for an industrial than it is for a utility. It may often be neglected without affecting results significantly.

Mr. Backenstoss cites advantages in taking the income-statement approach rather than the tax-form approach: (1) it gives a better measure of a company's profitability, and (2) it avoids thinking that the income tax rate should be 52%. It may be true that over-all company operations may be more truly shown by an income-statement approach which differs from a tax-form approach. Nevertheless, for an incremental economic analysis, the tax-form approach gives correct results, because it correctly measures the cash flows, as defined by equation 16 in the paper, from which the rate of return on the investment is derived.

Mr. Backenstoss infers that the paper suggests the use of 52% as the income-tax rate. He believes that this is an error since the actual rate is much lower because of accelerated tax-allowable depreciation.

The author believes that the correct tax rate to use in an economic analysis is not the effective tax rate but the maximum rate applicable on incremental pretax earnings. According to present income tax rules, this rate is 52% for taxable earnings in excess of $25,000. For the larger industrial companies, the maximum rate of 52% would usually be the one used. Most writers on problems involving economic analysis seem to agree with this view because they use 50 or 52% as a tax rate in the examples they cite. Furthermore, in Barrons, August 5, 1957, figures are given for U.S. corporate profits after tax and before tax during the past year. These have a ratio of 0.49, or an indicated average tax rate of 51%.

The figure of 15% for an average rate of return is questioned by Mr. Backenstoss. The rate which was suggested in the paper for economic comparisons was not this, but a rate determined by company policy, or by finding the company average return on their net assets, and adjusting this to the risk of the proposed investment as compared with the average risk. It may also be adjusted to allow for increases in prices, as Mr. Backenstoss feels is required, although the author believes, in view of very substantial expansion programs carried out over the past few years by many industrial companies, at prices close to today's, that this factor is of less importance than Mr. Backenstoss seems to indicate.

As Mr. Backenstoss points out, the comparison of two alternatives using differential earnings presupposes that both alternatives have the same investment charge factor. However, the equation $I=E/R$, as developed in the paper, can be used to compare two alternatives having different investment charge factors. This is done by comparing each with a third alternative of no investment. After computing gross earnings E for each plan, the justifiable investments for each may be found and compared with each other and with the base case of making no investment.

REFERENCES

1. FORCED OUTAGE RATES OF HIGH-PRESSURE STEAM TURBINES AND BOILERS, AIEE Joint Subcommittee Report. *AIEE Transactions*, vol. 76, pt. III, June 1957, pp. 338–43.

2. SOME NEW MATHEMATICAL ASPECTS OF FIXED CHARGES, Constantine W. Bary, W. T. Brown. *Ibid.*, pp. 230–49.

Application of a Mechanical Rectifier to a Cyclotron

JACK WARREN
ASSOCIATE MEMBER AIEE

THE first mechanical rectifier to be used in the field of "atom-smashing" has been installed and has completed 8 months of operation. Anticipated high reliability and rectification efficiency has been realized. To date no major operating or maintenance problem has arisen.

In February 1956, a 3,000-ampere 450-volt d-c mechanical rectifier manufactured by the I-T-E Circuit Breaker Company was placed in service at Columbia University's 385 million electron-volt synchrocyclotron at Irvington, N. Y., as the power supply for the cyclotron electromagnet.

Reasons for Using Mechanical Rectifier

In 1953, although the cyclotron had been in operation since 1950, it was decided to obtain a new magnet power supply because of the poor performance record of the motor-generator set then in use, and because cyclotron revisions were being considered which would eventually require a power supply of higher rating.

The motor-generator set consisted of a 900-horsepower 2,300-volt 3-phase synchronous motor, and a 600-kw 300-volt shunt generator. During its 3 years of use, five major breakdowns caused a total loss of $13^1/_2$ months of cyclotron operating time.

Various manufacturers were invited to suggest a new power supply. Proposals were received for motor-generator sets, ignitrons, a selenium rectifier, and a mechanical rectifier. These devices were compared on the basis of initial cost efficiency, space requirements, cooling requirements, and reliability. In regard to the last requirement, a reliable device meant one that would break down only infrequently and when it did, could be repaired either by the manufacturer or by the cyclotron operating staff in a reasonably short time. Investigation indicated that any of these devices, if well designed and constructed, should be reliable. The reliability of the mechanical rectifier was particularly investigated since this was the newest device. Every user of an I-T-E mechanical rectifier was questioned, and the findings were favorable. In comparing the efficiency of the various devices, consideration was given to fractional load efficiency, since at least the first few years of operation would be at about 40% of the power supply's rated load. On the basis of all the aforementioned considerations, the decision to purchase a mechanical rectifier was made.

Description of Cyclotron Magnet

The steel yoke of the magnet weighs 2,000 tons. It is 33 feet long and 21 feet high. The upper and lower coils for the magnet are each constructed of six spirally wound pancakes of copper bus bar. Each pancake is made up of 5,700 feet of 4-inch-wide by 9/32-inch-thick bus bar. The total copper weight of both coils is 150 tons. Each coil is in a 22-foot-outside-diameter coil tank which resembles a doughnut with a square cross section. These tanks are primarily liquid containers for the cooling oil.

The resistance of the magnet varies from 0.120 to 0.147 ohm, depending upon its temperature. The inductance varies from about 2 henrys to 30 henrys, depending upon the current. The total energy stored in the magnet at 3,000 amperes ($\int_0^{3,000} Li\,di$) is about 20,000,000 joules. This high inductance and stored energy were factors which weighed heavily in the design of the rectifier. If, while the magnet were carrying current, its input should open for any reason, dangerous overvoltages would be generated by the magnet. Damages to the magnet could be so severe that repair time, and hence cyclotron shutdown time, could be of the order of a year. Repair costs in case of severe damage could amount to the order of several hundred thousand dollars.

Description of Rectifier Equipment

Fig. 1 shows the schematic diagram for the basic rectifier circuit.

OUTDOOR SUBSTATION

The outdoor substation of the laboratory was expanded to accommodate the a-c switchgear and step-voltage regulator; see Fig. 2. The a-c switchgear unit is of weatherproof construction and contains the main a-c circuit breaker, rated at 1,200 amperes, 3 phase, 15,000 volts, with an interrupting capacity of 500 megavolt-amperes.

Other major components of the outdoor switchgear unit are an auxiliary power transformer, a battery rack and charger and potential and current transformers.

The battery consists of a bank of nickel-cadmium alkaline batteries which supplies 48 volts d-c for tripping power for the main a-c and d-c circuit breakers. The batteries were chosen for a reliable source of tripping power in case of a failure of the one 13.8-kv feeder. An automatic charger keeps these batteries charged. The potential and current transformers are for meters, instruments, and protective relay purposes.

The step-voltage regulator is rated at 239 kva, 13,800 volts, 3 phase, 60 cycles, 16% raise and lower. It is essentially a 3-phase regulating autotransformer of the oil-immersed self-cooling type. It is provided with taps in a regulating (series) winding and a multiposition, electrically operated tap-selector switch which will give automatically a range of voltages in the series winding. Thus, voltage regulation from 16% above to 16% below line voltage is obtained in 32 1% steps.

It should be noted that for most applications of the mechanical rectifier, a step-voltage regulator would not be necessary. The voltage regulation available from the taps on a normal rectifier transformer (about 10%) plus the voltage regulation available from the normal commutating reactors (approximately 50%) is usually ample enough. The reason that a step-voltage regulator was used in this application is that it was desired for a particular rectifier transformer tap to energize the cyclotron magnet at a minimum direct voltage to keep the rate of rise of magnet flux at a minimum. This would minimize the eddy currents developed in the copper dee structures, which in turn would limit undesirable physical stresses due to reaction between the flux set up by the eddy currents and the main flux of the magnet.

CYCLOTRON BUILDING

Fig. 3 shows the layout of the rectifier equipment in the cyclotron building.

The output of the step-voltage regulator in the outdoor substation is fed into the

---

Paper 57-700, recommended by the AIEE Industrial Power Rectifiers Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted September 17, 1956; made available for printing May 15, 1957.

JACK WARREN is with Singmaster and Breyer, New York, N. Y., and the Columbia University Nevis Cyclotron Laboratories, Irvington, N. Y.

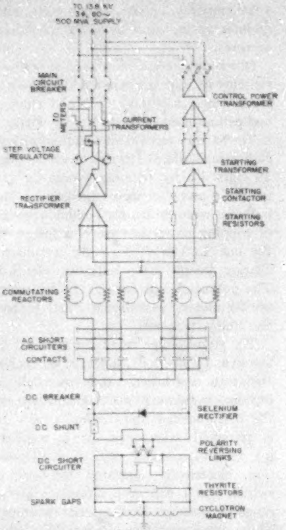


Fig. 1. Basic rectifier circuit, schematic diagram

cyclotron building to the rectifier transformer. The transformer is rated at 1700 kva, 15,900 volts, 3 phase, 60 cycles, 55-degree-centigrade temperature rise at full load, 3.4% impedance. It is Askarel-filled and self-cooled. There are five no-load taps which permits a total reduction of about 11% from maximum voltage. The no-load tap changer is controlled by an operating handle on the transformer tank. It is key interlocked to prevent its operation during load conditions.

The six commutating reactors are mounted in the same tank with the transformer winding. There are three control windings on these reactors in addition to the main load current winding: the break pre-excitation, the flux reversal, and the overlap.

The break pre-excitation winding supplies the reactor core with the current necessary to reverse its flux during the break step, thereby relieving the load winding from this duty and enabling the contact to open while it is carrying close to zero current. This winding, in conjunction with an arc suppressor, an *R-C* circuit across the contact, completes the flux change in the reactor after the contact opens so that the contact has time to separate sufficiently before it sees the inverse voltage.

The flux reversal winding enables the commutating reactor to be used for regulating the rectifier output voltage. This winding, together with the control circuit that feeds it, permits the flux of the reactor core to reverse a controllable amount before the closing of the contact, thereby enabling the time for completion of flux reversal after the closing of the contact, the make step, to be shortened a predictable amount of time. The controlling of the amount of flux reversal prior to a contact closing, and therefore the length of the make step after the contact closes, is similar to the use of phase-shift control of the firing point of grid-controlled mercury-arc rectifiers and ignitron rectifiers. The longer the make step, or phase shift delay, the less will be the output voltage.

The flux reversal windings, besides enabling the commutating reactors to regulate voltage over a wide range, enable the commutating reactors to control cyclotron current to a fine degree. For the sake of cyclotron beam stability, specifications called for the magnet power supply to regulate magnet current automatically to at least ±0.25% of a preset value.

The major components of the automatic current controller are shown in the block diagram in Fig. 4.

The cyclotron magnet current is sensed by a 3,000-ampere 30-millivolt shunt; it has a limit of error of 0.04%. The shunt output is fed to a self-balancing-type potentiometer current recorder-controller with a range span of 600 amperes. Four suppressed zero ranges, 1,000, 1,500, 2,000, and 2,500 amperes, are provided so that current on any one range can be read to about 1 ampere. The accuracy of the instrument is 0.1%× suppression ratio +0.3% of range span. The dead band of the instrument is 0.15% of the range span, or 0.9 ampere.

A second slidewire, in addition to that of the self-balancing potentiometer, is provided for control purposes. If the magnet current, as indicated by a black pointer on the instrument scale, deviates from a preset value as selected by a red pointer on the same scale, the sliding contact of the control slidewire moves off center an amount proportional to the deviation.

The position of the control slidewire is converted to current from zero to 5 milliamperes, by an electronic control unit. Proportional band, reset, and rate circuits are provided for adjusting for optimum automatic magnet current control.

The load in which this control unit develops the current is the signal winding of a magnetic amplifier. The d-c output of this magnetic amplifier, the flux reversal current, is fed to another control circuit which feeds the flux reversal winding of the commutating reactor. A decrease of magnet current causes the automatic current regulator system to increase the amount of flux reversal in the commutating reactor core before the contact closes, thereby decreasing the length of the make step, and thereby increasing the rectifier output voltage and hence the cyclotron magnet current.

The third control winding of the commutator reactor is the overlap winding. Overlap time is the time that the contact about to give up the load current and the contact about to take it are both closed. An inherent characteristic of this type of rectifier is that to keep contacts breaking in the middle of their break step, the overlap time must vary with input voltage, phase control, and load. Automatic control of the overlap time, while maintaining the same closing time for contacts, is accomplished by the overlap regulator and circuit. Position of the beak, during the break step, is fed to the overlap regulator control circuit via the overlap winding on the commutating reactor.

High-speed a-c and d-c short-circuiting breakers are provided to protect contacts in case of backfire. These are latched against spring tension in the open position by the flux of a permanent magnet. Tripping is caused by shifting flux by energizing a coil wound on the magnet. This coil is energized by impulse transformers when these sense unbalances between the a-c and d-c systems of the rectifier. The breakers close within about a millisecond of the initiation of the trigger pulse. The d-c short circuiter is also one of several devices for insuring that the cyclotron magnet has a discharge path in case of trouble.

A single-pole 3,000-ampere d-c circuit breaker is in one of the output busses of the rectifier. It is apparent that this breaker cannot by itself open the highly inductive load, as the maximum stored energy of the magnet is approximately 20,000,000 joules and the maximum arc energy that could be dissipated by any d-c breaker is under 1,000,000 joules. And as previously stated interrupting the magnet circuit while it was carrying current would cause the magnet to generate dangerous overvoltages.

Use of the d-c breaker was made possible by a bank of selenium rectifiers located in the breaker cubicle. The bank is connected in reverse polarity as shown in Fig. 1, so that it does not conduct under normal operation. When the d-c breaker is opened, the voltage at the terminals of the load reverses and the energy stored in

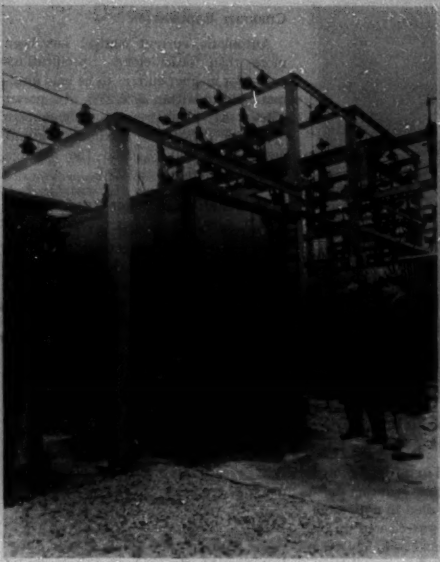


Fig. 2. Substation showing step-voltage regulator at left, a-c switchgear in center

Fig. 3. Mechanical rectifier equipment in cyclotron building

the magnet dissipates itself in the resistances of the magnet and selenium rectifiers. The normal leakage power dissipation at 450 volts will be about 300 watts or 0.23% of rated.

To protect the expensive cyclotron magnet further against any unforeseen contingency, a bank of Thyrite resistors is connected across the copper bus to the magnet. The Thyrite is rated to dissipate the maximum energy of 20,000,000 joules while limiting the maximum induced voltage across the magnet terminals to 1,700 volts. Normal dissipation in the Thyrite at the 450-volt level will be about 3 kw, or 0.23% of rated rectifier load. At the present operating level of 260 volts, the dissipation is only about 100 watts, or 0.02% of the rectifier output.

To protect the magnet still further, a pair of spark gaps is installed at each of the magnet coil terminals to ground. A spark gap consists of two $4\frac{3}{4}$-inch-diameter copper disks, 1/2-inch thick, separated by 0.015 inch of mica, all enclosed in a micarta enclosure. Tests have shown that these gaps break down at about 1,700 volts.

It is necessary, at times, to reverse the polarity of the current through the magnet. This is done by first shutting down the rectifier and permitting the magnet current to decay to a low level. Then, when the residual current is low, a special double-pole double-throw "make-before-break" polarity-reversing switch is operated. This switch is located in the d-c breaker cubicle. The door of the cubicle is key interlocked so that the switch cannot be reached unless the d-c breaker is open. A meter at the switch indicates the amount of residual magnet current.

After the switch is operated, the selenium presents inverse resistance to the residual current. Depending upon the value of this inverse resistance and the value of the residual current, a high voltage could be induced by the magnet. To eliminate this possibility, a 300-ohm resistor is permanently connected across the selenium.

CONTROLS

Startup is initated by the starting control switch. This energizes the rectifier mechanism through the starting transformer and starting resistors. The main a-c breaker control switch is then operated which energizes the rectifier through the main transformer. The next step is to energize the magnet by operating the d-c breaker control switch and closing the breaker. This breaker is interlocked so that it cannot be closed unless both the step-voltage regulator and phase control

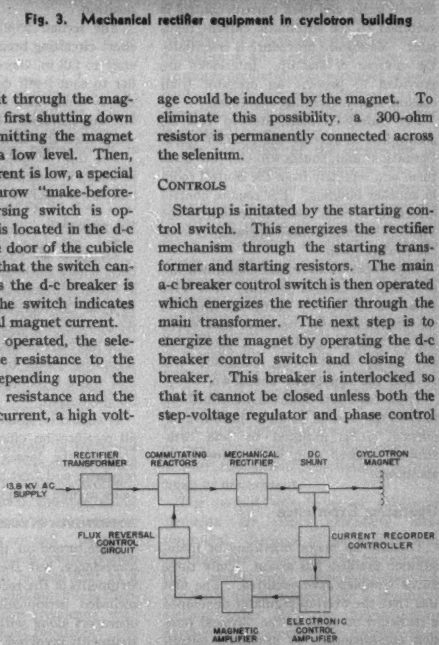


Fig. 4. Automatic current control; block diagram

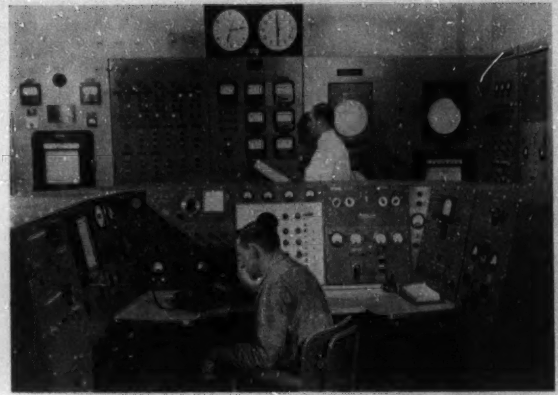


Fig. 5. Cyclotron control room with rectifier control cubicle at upper left

are in the minimum voltage position. As previously discussed, it is desired to energize the magnet at minimum voltage so as to keep eddy currents in the dee structures to a minimum. Magnet current is then adjusted manually to the desired value and then current control is switched to the automatic current regulater. Shutdown procedure is essentially the reverse of startup. Interlocks are provided to prevent an operator from causing faulty operating sequences.

For safety reasons, startup can be accomplished only at the rectifier itself. Operation and shutdown can be done there and remotely from the control room in another building. For remote control a rectifier control board is provided; see Fig. 5. Normal rectifier operation is from the control room.

Two types of automatic shutdown are provided in case of trouble. The first type, the emergency shutdown, drops in the short-circuiting breakers and opens the a-c and d-c breakers. A backfire or any other short-circuit condition will initiate the emergency shutdown.

The second type of automatic shutdown follows the normal manual shutdown sequence. It is initiated by less catastrophic causes than is the emergency shutdown.

Operating Experience

To date no major operating or maintenance problem has arisen. Only three minor troubles have occurred. The first was that the overlap regulator developed a tendency to stick in the physical position corresponding to minimum output voltage. This resulted in a backfire, and the destruction of one contact during shutdown, if the operator failed to notice the sticking. The trouble was cleared up by adding a spring which acts to pull the regulator out of the minimum voltage position.

The second trouble was that the a-c short-circuiting breaker developed a tendency to fall in, thereby causing the rectifier to shut itself down. The cause was simply some dirt between the magnetic latching surfaces. Cleaning the surfaces eliminated the trouble.

The third trouble was a short-circuited insulator in the current recorder-controller which was easily repaired.

EFFICIENCY

To date, all cyclotron operation has been with a 2,000-ampere magnet current at about 260 volts corresponding to about 40% of the rectifier's rated power load. At this load the efficiency is about 94.7%. This has been determined by using an a-c watt-hour meter to determine a-c input, and the d-c recorder and d-c voltmeter for d-c output.

It should be noted that this is an overall rectification efficiency and takes into account the losses in the step-voltage regulator and that most rectifier installations require no step-voltage regulator.

ATTENDANCE NECESSARY

Once turned on, the rectifier requires no attendance, but the oscilloscope and instruments of the rectifier are checked and recorded periodically by the cyclotron operators along with the many other instruments involved in the operation.

CURRENT REGULATION

Automatic current control has been more than satisfactory. Specifications called for magnet current to be held automatically to within ± 0.25% of a preset value. At the present 2,000-ampere operating point, the maximum current deviation is ±0.1%, and the average current deviation is ±0.05%.

PROCEDURE FOR CONTACT REPLACEMENT

To replace a complete set of six contacts and return the rectifier to service takes about 20 minutes and requires no special skill. Timing of the new contacts is checked by running the mechanism in the "test" position and checking a voltmeter for a predetermined reading. Little, if any, timing adjustment is ever required. Timing adjustment is accomplished by rotating an external screw geared to each individual push rod.

With the contacts replaced and the timing checked, the rectifier is ready to deliver load; no "break-in" period is required.

CONTACT LIFE

To date contact life has been satisfactory. At the end of 613 hours of operation, the six contacts were removed and inspected for an estimate of contact life. The estimate was 2,000 hours.

Only one backfire has occurred during the cyclotron operation. It was caused by a line voltage disturbance caused by the failure of another feeder in the power company's distribution system. All contacts were damaged and replaced.

Several other contacts have been damaged and replaced due to backfires occurring during the shutdown sequence when operators failed to notice that the overlap regulator was sticking in the minimum voltage position. As previously stated, this trouble has been eliminated.

One contact was replaced after 852 hours of operation when observation of its break via the oscilloscope indicated that it was breaking erratically.

With the present 2,300 operating hours on the rectifier, five of the present contacts have been in use for 1,600 hours and the sixth for 850 hours.

Conclusions

The mechanical rectifier, after 8 months of operation, is proving to be a reliable and efficient power supply for a cyclotron magnet. Automatic current control, to a fine degree, has been obtained. The inherent dangers, due to the high stored energy in the inductance of the magnet, have been surmounted.

Discussion

L. K. Dortort (I-T-E Circuit Breaker Company, Philadelphia, Pa): Mr. Warren has given us a comprehensive and detailed report on the preliminary investigations and selection of a mechanical rectifier for a very special application, the engineering characteristics developed to make the application practical, and the operating experience with this mechanical rectifier supplying power to the magnet coil of a cyclotron having an extremely large stored energy. It is gratifying to have the engineering decisions and calculations of the manfacturer so accurately confirmed.

One of the most difficult and complex problems, and the one involving the greatest risk, involved the starting, stopping, and reversing of current flow from a mechanical rectifier to an extremely inductive load. Specifications restricted the maximum voltage appearing across the cyclotron coil for any reason including inductive voltages, to a value of 500 volts, only 50 volts above rated voltage. A similar restriction was placed on maximum current. These stringent requirements were well justified by the high cost of any repairs to the magnet coil in time and money.

Adding to the difficulty was the desire to shut down, reverse polarity, and start up again in 30 minutes. The natural decay time of the magnet current from 3,000 amperes down to 50 was calculated to be approximately 550 seconds. Considering the fact that below 50 amperes the inductance of the cyclotron coil is above 36 henrys, it was obvious that no switching could be safely attempted at coil currents exceeding a small fraction of an ampere. Obviously, switching arrangements could be devised to insert increasing resistance values as the current decreased, but aside from their complexity such switching arrangements involved a considerable risk of malfunctioning and damage to the cyclotron.

With the scheme finally adopted the estimated time for the decay of current from 3,000 amperes down to 0.6 ampere is $17^{1}/_{2}$ minutes and does not require any switching within the coil circuit. Special selenium rectifier stacks were designed and connected across the coil as a voltage-suppressor valve. The characteristics of the selenium valve were chosen so that the rate of decay of flux in the magnet, interacting with the eddy currents in the conducting parts within the *D*'s, would not produce excessive forces. Aluminum base plates were chosen for low reactance and the physical design balanced for thermal capacity. At 0.6 ampere, it is safe to use the make-before-break reversing switch, and normal startup in the reverse direction can be initiated.

An Analytic Method for Finding the Closed-Loop Frequency Response of Nonlinear Feedback-Control Systems

KATSUHIKO OGATA
ASSOCIATE MEMBER AIEE

THE linear theory in the field of feedback-control systems is highly developed and various techniques are available for analysis and synthesis of linear feedback-control systems. On the other hand, there are many problems to be solved in the field of nonlinear feedback-control systems. Techniques of analyzing and synthesizing nonlinear feedback-control systems are quite limited because of difficulties in handling nonlinear differential equations.

It is well known that frequency-response curves of feedback-control systems with saturation nonlinearities often exhibit the jump-resonance phenomenon. It is also known that responses of nonlinear feedback-control systems sometimes exhibit subharmonic oscillations.

The jump-resonance phenomenon is discussed or explained by authors such as Minorsky,[1] Levinson,[2] Prince, Jr.,[3] and Truxal.[4] Levinson,[2] using an analytical and graphical technique, indicates the multivalued response and the jump-resonance phenomenon of a tachometer-stabilized system. The principle used in his paper is that the effect of saturation is to lower the average-gain value of the path in which saturation takes place. Prince, Jr., using the equivalent-gain method, shows in his paper[3] the jump resonance phenomenon as responses of an acceleration-limited system and an error-limited system.

Subharmonic oscillations as responses of nonlinear vibratory systems are often discussed in the field of nonlinear mechanics. Not much is known, however, about subharmonic oscillations as responses of nonlinear feedback-control systems. When a nonlinear feedback-control system includes saturation, it is seen that at a fixed input frequency and amplitude there may be more than one frequency with which the system output oscillates. For example, when a nonlinear high-gain feedback-control system with saturation is used as a recorder, the output of the recorder may oscillate with a frequency which is $1/n$ (n is an integer greater than one) times the input frequency, if the input of the recorder is periodic with a high frequency. When the frequency of the output oscillation is $1/n$ times that of the input, the output oscillation may be called the $1/n$-order pure subharmonic oscillation. An analysis of pure subharmonic oscillations as responses of feedback-control systems with saturation nonlinearities will be presented by the author in another paper.

The possibility of the occurrence of pure subharmonic oscillations as the response of a nonlinear feedback-control system can be predicted from an accurate closed-loop attenuation-frequency curve for a given input amplitude. For example, when a nonlinear feedback-control system includes saturation, and if the maximum ratio of the system output amplitude to the system input amplitude is large (about 3 decibels or more), there is a possibility that the output of the system may oscillate with a subharmonic of the input frequency. However, if the maximum ratio of the output amplitude to the input amplitude is small (2 decibels or less), the possibility of the occurrence of pure subharmonic oscillations is very small or nonexistent.

The output waveforms of nonlinear elements are often extremely nonsinusoidal. Because of the difficulty in handling such nonsinusoidal waveforms, the well-known describing-function method[5,6] which is essentially a graphical method takes only the fundamental-harmonic component into consideration

---

Paper 57-634, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted October 23, 1956; made available for printing April 16, 1957.

KATSUHIKO OGATA is with the University of Minnesota, Minneapolis, Minn.

The material in this paper is based on a part of a dissertation submitted by the author in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Engineering Science at the University of California. The author wishes to acknowledge with sincere gratitude many discussions offered during the course of the research by Prof. A. M. Hopkin who acted as chairman of the Dissertation Committee.

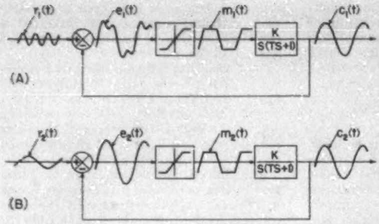


Fig. 1. Block diagrams of second-order feedback-control system with saturation nonlinearity showing

A—Pure subharmonic oscillation
B—Fundamental oscillation

for analyzing and synthesizing nonlinear feedback-control systems. To the author's knowledge, no analytic methods for finding closed-loop frequency-response characteristics of nonlinear feedback-control systems and/or jump-resonance frequencies are found in past publications.

The main purpose of this paper is to present an analytic method for finding the accurate closed-loop frequency-response characteristics of nonlinear feedback-control systems. By means of the periodic-response function, defined herein as the Laplace transform of the steady-state waveform of the output of a nonlinear element when the input to that element is periodic, this analytic method takes all or most of the harmonic components of the output of the nonlinear element into consideration for finding the closed-loop frequency-response characteristics. Accurate closed-loop frequency-response results obtained by this method may be used for predicting the possibility of the occurrence of pure subharmonic oscillations. In this paper methods are also presented for finding the jump-resonance frequencies in terms of the system constants and the input amplitude.

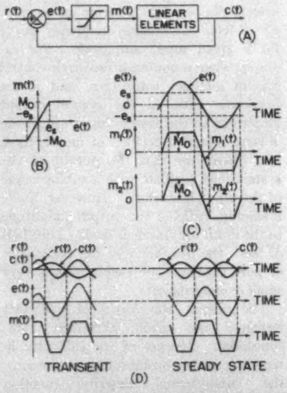


Fig. 2. Block diagram, nonlinearity, waveforms c(t), e(t), and m(t) of nonlinear feedback-control system

The emphasis of this paper is placed on showing the general approach of the method. Therefore, nonlinear second-order feedback-control systems are used as examples. However, the method presented is applicable to nonlinear higher order feedback-control systems. An application of this method for finding the closed-loop frequency-response characteristics of a nonlinear higher order feedback-control system has been shown previously.[7]

Differences Between the Method Presented and the Describing-Function Method

Some differences between the method presented in this paper and the describing-function method are listed in the following.

1. The author's method takes into account the effects of all or most of the higher harmonics of the nonsinusoidal output of the nonlinear element for finding the system output, while the describing-function method takes into consideration only the fundamental harmonic of the nonsinusoidal output of the nonlinear element.

2. The author's method is essentially analytical, while the describing-function method is essentially graphical.

3. The author's method does not require the assumption that the input to the nonlinear element is sinusoidal. On the contrary, it is the fundamental assumption of the describing-function method that the input to the nonlinear element is sinusoidal; therefore, the method presented in this paper is valid for finding closed-loop periodic-response characteristics of nonlinear feedback-control systems for periodic inputs such as sinusoidal inputs, square-wave inputs, triangular inputs, etc., while the describing-function method is valid only for sinusoidal inputs.

4. The author's method is valid for finding the pure subharmonic output of the system. Fig. 1(A) shows a pure subharmonic oscillation of the output of a feedback-control system with a saturation nonlinearity. Fig. 1(B) shows a fundamental oscillation of the output of the same system. The steady-state waveform of $e_1(t)$ is quite different from a sinusoidal wave, therefore, the describing-function method is not valid when the system output exhibits the pure subharmonic oscillation as shown in Fig. 1(A). However, the method presented in this paper is valid, because the waveform of $m_1(t)$ is used for finding $c_1(t)$ analytically.

Analytic Method for Finding Closed-Loop Frequency Response of Nonlinear Feedback-Control Systems

PERIODIC-RESPONSE FUNCTION

Fig. 2(A) shows the block diagram of a feedback-control system with a saturation nonlinearity. The input-output characteristic of the nonlinear element is shown in Fig. 2(B). The exact output waveform of the nonlinear element may be shown as the curve $m_1(t)$ in Fig. 2(C). The waveform $m_2(t)$ is a straight-line approximation of $m_1(t)$. The error signal $e(t)$ and the waveform $m_2(t)$ are also shown in Fig. 2(C). The approximate waveform $m_2(t)$ may be the same as the exact waveform $m_1(t)$, when the latter has only straight-line segments, e.g., the output waveform of an on-off nonlinearity. After a sinusoidal input is applied to the system and steady state has been reached, the output of the nonlinear element will be periodic. Let the output waveform of the nonlinear element be $m(t)$. Fig. 2(D) shows the transient and steady-state waveforms $e(t)$, $m(t)$, and $c(t)$ together with the input $r(t)$.

By defining $M(s)$ which is the Laplace transform of $m(t)$,the steady-state waveform of the output of the nonlinear element when the input to that element is periodic, as the "periodic-response function," it may be written as follows,[8] if the period of the waveform $m(t)$ is $2a$.

$$M(s)=\frac{\int_0^{2a}e^{-st}m(t)dt}{1-\epsilon^{-2as}} \qquad (1)$$

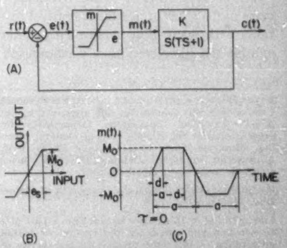


Fig. 3. A—Block diagram of second-order feedback-control system with saturation nonlinearity. B—Input-output characteristic of nonlinearity. C—Waveform m(t)

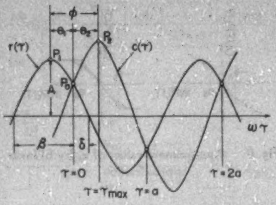


Fig. 4. Definition of phase angle ϕ

Example: Periodic-Response Function of Saturation Nonlinearity. Fig. 3(A) shows the block diagram of a second-order feedback-control system with a saturation nonlinearity. The input-output characteristic of the nonlinearity is shown in Fig. 3(B). The steady-state waveform of the signal $m(t)$ is shown in Fig. 3(C). When the degree of saturation is relatively high, the waveform of the output of the saturation nonlinearity is trapezoidal. The periodic-response function of the saturation nonlinearity may be obtained as a function of M_o, a, and d, if M_o and d are defined as follows: M_o is the maximum output of the nonlinearity, and d is the time interval when the output changes from zero to the point where saturation starts as shown in Fig. 3(C). In general, d is a function of a, the half-period of the output, and E, the amplitude of the error signal. Using equation 1, the periodic-response function of the saturation nonlinearity, the Laplace transform of the waveform $m(t)$ shown in Fig. 3(C) may be written as follows

$$M(s)=\frac{M_o}{ds^2}\times \left[1-(\epsilon^{-ds}+\epsilon^{ds-as})\sum_{n=0}^{\infty}(-1)^n\epsilon^{-ans}\right] \tag{2}$$

METHOD OF APPROACH

The analytic method presented in this paper uses the following procedure:

1. Find the exact or the approximate waveform $m(t)$, the steady-state output of a nonlinear element.
2. Find $V(s)$, the Laplace transform of the system output velocity $v(t)$, in terms of the periodic-response function $M(s)$, the Laplace transform of $m(t)$, and the system transfer functions.
3. Find the inverse Laplace transform of $V(s)$ and obtain the output velocity as a function of the output frequency.
4. Integrate the output velocity with respect to time and find the system output $c(t)$ as a function of the output frequency. Obtain the relation between the output amplitude and the output frequency; also obtain the relation between the output phase lag and the output frequency.

CLOSED-LOOP ATTENUATION-FREQUENCY CHARACTERISTICS

When the system input $r(t)$ is sinusoidal i.e., $r(t)=A\sin(\omega t+\alpha)$, the attenuation-frequency characteristics for the second-order feedback-control system with a saturation nonlinearity shown in Fig. 3(A) may be found by obtaining the equation expressing the ratio of the output amplitude C_p to the input amplitude A, as a function of the output frequency. In Fig. 3(A), $c(t)$ and $m(t)$ are related by the following equation

$$(Tp^2+p)c(t)=Km(t) \tag{3}$$

where $p=d/dt$. Taking the time origin $t=0$ so that $r(0)=c(0)$, $e(0<t<a)>0$, and using equations 2 and 3, $V(s)$, the Laplace transform of $v(t)$ where $v(t)=pc(t)$, may be written as follows

$$V(s)=\frac{KM_0}{d}\left(\frac{1}{s^2}-\frac{T}{s}+\frac{T}{s+\frac{1}{T}}\right)\times \left[1-(\epsilon^{-ds}+\epsilon^{ds-as})\sum_{n=0}^{\infty}(-1)^n\epsilon^{-ans}\right]+\frac{v(0)}{s+\frac{1}{T}} \tag{4}$$

where $v(0)$ is the initial velocity. Equation 4 shows the Laplace transform of the velocity $v(t)$ of the system output when the input to the linear element, whose transfer function is $K/s(Ts+1)$, has a trapezoidal waveform as shown in Fig. 3(C).

It should be noted here that in the case of the nonlinear second-order feedback-control system, the classical method to solve equation 3 in time domain may be simpler. The purpose of using the Laplace-transform approach shown here and in the appendix is to describe the procedure equally applicable to nonlinear higher order feedback-control systems. When the linear elements of the nonlinear feedback-control system have double or triple poles, it will be difficult to find periodic solutions by working in the time domain.

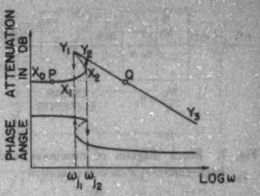


Fig. 5. Attenuation-frequency and phase-frequency curves showing jump resonance

Taking the inverse Laplace transform of equation 4 for the time interval $\lambda a \leqslant t \leqslant (\lambda+1)a$, where a is the half-period of the output and λ is a sufficiently large odd number, letting $t-\lambda a=\tau$, and integrating $v(\tau)$ with respect to τ, the system output $c(\tau)$ may be obtained as a function of a and τ. If the velocity of the system output becomes zero at some instant during the time interval d and $a-d$, shown in Fig. 3(C), the maximum of $c(\tau)$ for a half-cycle $(0\leqslant\tau\leqslant a)$ occurs during the time interval $d\leqslant\tau\leqslant a-d$. Let $\tau_{\max}$ be the time when the output position during a half-cycle $(0\leqslant\tau\leqslant a)$ becomes maximum. Integrating $v(d\leqslant\tau\leqslant a-d)$ with respect to τ and substituting $\tau=\tau_{\max}$, the output amplitude C_p may be obtained as follows:

$$C_p=KM_oT\left(\ln\cosh\frac{a}{2T}-\ln\sinh\frac{d}{T}+\ln\frac{d}{T}\right) \tag{5}$$

Equation 5 is the analytic solution for finding the system output amplitude C_p of the system shown in Fig. 3(A) as a function of a and d. In equation 5 d is a

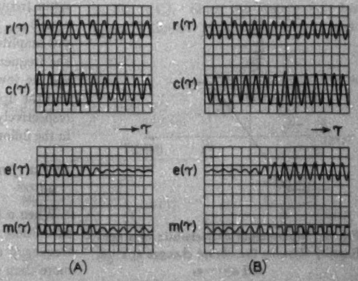


Fig. 6. Curves $r(\tau)$, $c(\tau)$, $e(\tau)$, and $m(\tau)$

A—Showing lower jump resonance

B—Showing upper jump resonance

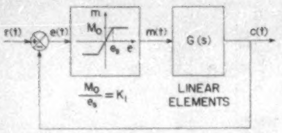


Fig. 7. Block diagram of nonlinear feedback-control system with saturation

function of a and E, the amplitude of the error signal, which is a function of the amplitude and frequency of the input. Substituting $a=\pi/\omega$ in equation 5 and dividing both sides of the equation by A, the following equation may be obtained:

$$\frac{C_p}{A}=\frac{KM_oT}{A}\times\left(\ln\cosh\frac{\pi}{2T\omega}-\ln\sinh\frac{d}{T}+\ln\frac{d}{T}\right) \quad (6)$$

The attenuation-frequency curve of the system shown in Fig. 3(A) under nonlinear operation may be obtained from equation 6.

Closed-Loop Phase-Frequency Characteristics

Since the waveform of the system output is not sinusoidal for nonlinear operation, the phase difference between the input and the output is defined in this paper as the phase difference between the successive maxima of the system input and the system output.

The phase lag ϕ is shown in Fig. 4. In Fig. 4 the input is sinusoidal, i.e., $r(\tau)=A\sin(\omega\tau+\beta)$, and β is the initial phase angle of the input with respect to τ. As may be seen from Fig. 4, the phase lag ϕ is the sum of θ_1 and θ_2, where θ_1 is the phase difference between P_0, the point which corresponds to $\tau=0$, and P_1, the point which corresponds to the maximum of the system input $r(\tau)$; and θ_2 is the phase difference between P_0 and P_2, the point which corresponds to the maximum of the system output $c(\tau)$ for a particular output frequency. Since $c(0)$, the system output magnitude at $\tau=0$, equals $A\sin\beta$; θ_1 is the difference between β and 90 degrees; and θ_2 equals $\omega\tau_{\max}$, the phase lag ϕ of the output of the system shown in Fig. 3(A) may be written as a function of ω, the system output frequency, and d as follows:

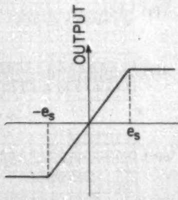


Fig. 8. Input-output characteristic of a nonlinearity showing the abrupt changes at $e=e_s$ and $e=-e_s$

$$\phi=180^\circ\left[1-\frac{T\omega}{\pi}\left(\ln\cosh\frac{\pi}{2T\omega}+\ln\frac{d}{T}-\ln\sinh\frac{d}{T}\right)\right]-\delta \quad (7)$$

where

$$\sin\delta=\frac{KM_oT}{Ad}\times\left[-T+T\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{\pi}{2T\omega}\right)+\frac{\left(\frac{\pi}{\omega}-d\right)d}{2T}\right],\ 0^\circ\leqslant\delta\leqslant 90^\circ \quad (8)$$

and the phase angle ϕ is obtained in degrees. Angle δ is indicated in Fig. 4.

Jump-Resonance Frequencies

The response of a nonlinear feedback-control system with saturation is multivalued when the system input lies within a certain range of amplitude and frequency. As shown in Fig. 5, if a sinusoidal input of a low frequency is fed to a nonlinear feedback-control system with saturation and this frequency is gradually increased, the point P on the attenuation-frequency curve representing the output amplitude of the system follows the curve $X_0X_1X_2$. When the input frequency equals ω_{j2}, the amplitude and the phase of the output change abruptly as shown in Fig. 5. With a further increase in the input frequency, the point Q on the attenuation-frequency curve follows the curve Y_2Y_3. If the procedure is reversed and the input frequency is decreased gradually from a higher value, the point Q follows the curve $Y_3Y_2Y_1$. When the input frequency is decreased further and equals ω_{j1}, the abrupt changes in the output amplitude and phase occur. Calling the frequencies corresponding to ω_{j1} and ω_{j2} the lower jump-resonance frequency and the upper jump-resonance frequency respectively, they may be found as shown in the following.

Lower Jump-Resonance Frequency ω_{j1}

When a feedback-control system has a saturation nonlinearity, at a fixed input frequency and amplitude, there may be more than one $m(t)$ signal which satisfies the loop. In Fig. 5, when the frequency of the input lies between ω_{j1} and ω_{j2}, there can be two stable waveforms of $m(t)$ signal. These are a trapezoidal waveform corresponding to a nonlinear operation of the system and a sinusoidal waveform corresponding to a linear operation of the system. In Fig. 5 curve $X_0X_1X_2$ corresponds to a linear operation and curve $Y_1Y_2Y_3$ corresponds to a nonlinear operation of the system. When a system has a saturation nonlinearity, the magnitude of the system output is always limited. The lower jump-resonance frequency ω_{j1} corresponds to the largest output magnitude and is the lowest input frequency for which the system may be operated nonlinearly.

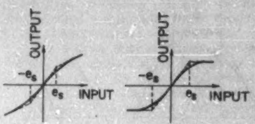


Fig. 9. Approximate values of e_s by bilinear approximation

At the intersection of curves $r(\tau)$ and $c(\tau)$ at $\tau=0$ in Fig. 4,

$$c(0)=A\sin\beta \quad (9)$$

Analytically, the output $c(0)$ of the system shown in the Fig. 3 may be written as a function of a, the half-period of the output, and d as follows:

$$c(0)=\frac{KM_oT}{d}\times\left[-T+T\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\times\tanh\frac{a}{2T}\right)\right]+\frac{KM_o}{2}(a-d) \quad (10)$$

The detail of the derivation of equation 10

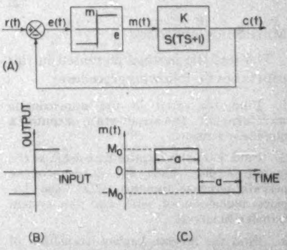


Fig. 10. A—Block diagram of second-order feedback-control system with on-off nonlinearity. B—Input-output characteristic of nonlinearity. C—Waveform m(t)

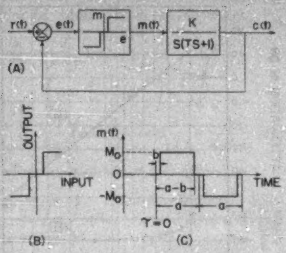


Fig. 11. A—Block diagram of second-order feedback-control system with on-off and dead-zone nonlinearity. B—Input-output characteristic of nonlinearity. C—Waveform m(t)

is shown in the appendix. For the system shown in Fig. 3, it can be operated nonlinearly for the region 90 degrees $<\beta<$ 180 degrees. Examination of equation 10 indicates that $c(0)$ becomes large as the value of a, the half-period of the output, increases, and vice versa. For the region $\beta<90$ degrees, the nonlinear operation is unstable. When β becomes slightly less than 90 degrees, the value of $c(0)$ becomes smaller than that of $c(0)$ corresponding to $\beta=90$ degrees, with the result that the amplitude and phase lag of the output become smaller than those of the output corresponding to $\beta=90$ degrees. The decrease in the amplitude and phase lag of the output makes the amplitude of the error signal smaller than that of the error signal corresponding to $\beta=90$ degrees. This causes the resulting amplitude of the output to be smaller than the preceding one. The phase lag of the output also becomes smaller than the preceding value. This process continues until the steady state is reached. For the system shown in Fig. 3, the steady state corresponds to the state where the system output oscillates in a linear fashion. The process may be seen from curve $c(\tau)$, the system output; curve $e(\tau)$, the error signal; and the curve $m(\tau)$, the output of the nonlinear element, as shown in Fig. 6(A). It is clear from equation 5 that the amplitude of the system output under nonlinear operation becomes large as the value of a, the half-period of the output, increases. Therefore, C_p becomes maximum when $c(0)$ becomes maximum. From equation 9, $c(0)$ becomes maximum when β equals 90 degrees. Let the value of a, the half-period of the output, which satisfies the equation

$$c(0)=A \tag{11}$$

be a_1. Then the lower jump-resonance frequency ω_{f1} may be obtained as the following equation

$$\omega_{f1}=\frac{\pi}{a_1} \tag{12}$$

Upper Jump-Resonance Frequency ω_{f2}

When the frequency of the input to the system with a saturation nonlinearity shown in Fig. 7 is low, the amplitude of the error signal is small, with the result that the system is operated linearly. When the frequency of the system input is increased from a low value, at a certain input frequency the amplitude of the error signal equals e_s, where e_s is the value of the input to the nonlinear element at which the input-output relation of the nonlinearity changes abruptly as shown in Fig. 8. When the amplitude of the error signal exceeds e_s, the magnitude of $m(\tau)$ (the output of the nonlinear element) does not increase as much as it should if the system were linear; the result is that the amplitude of the error signal for the following cycle becomes larger than before as does the phase lag of $c(\tau)$, the output of the system, for the following cycle, until the equilibrium state is reached. This equilibrium state corresponds to the nonlinear operation of the system and the waveform $m(\tau)$ is trapezoidal. The upper jump resonance may be seen from curve $c(\tau)$ shown in Fig. 6(B). Curves $e(\tau)$ and $m(\tau)$ at the upper jump resonance are also shown in Fig. 6(B). As may be seen from the curves in Fig. 6(B), the upper jump resonance takes place when the amplitude of the error signal equals e_s.

When there is no abrupt change in the input-output characteristic of the nonlinearity, the approximate value of e_s may be found by the bilinear approximation as shown in Fig. 9. In the cases of the nonlinearities shown in Fig. 9, the changes in the shapes of curves $e(\tau)$ and $m(\tau)$ take place gradually when the amplitude of the error signal $e(\tau)$ approaches e_s from a smaller value. When the amplitude of the error signal $e(\tau)$ approximately equals e_s, the upper jump resonance takes place.

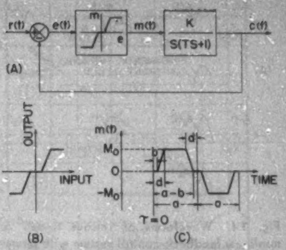


Fig. 12. A—Block diagram of second-order feedback-control system with saturation and dead-zone nonlinearity. B—Input-output characteristic of nonlinearity. C—Waveform m(t)

Since the limit of the linear operation of the system corresponds to the frequency at which the magnitude of the error signal equals e_s, the upper jump-resonance frequency ω_{f2} of the output of the system shown in Fig. 7 may be found to be frequency ω satisfying the following equation

$$\left|\frac{A}{1+K_1G(j\omega)}\right|=e_s \tag{13}$$

In equation 13, K_1 is the linear gain of the saturation nonlinearity, G is the transfer function of the linear elements, and A is the amplitude of the input sinusoid.

Analytic Frequency-Response Solutions of Nonlinear Second-Order Feedback-Control Systems

Second-Order Feedback-Control System with Saturation Nonlinearity

The analytic frequency-response solution of the system shown in Fig. 3(A) under nonlinear operation may be obtained from equations 6 through 8, if the maximum of the system output for a half-cycle ($0<\tau<a$) at steady state occurs during the time interval $d<\tau<a-d$, shown in Fig. 3(C). Since the system may be operated linearly if the input frequency is

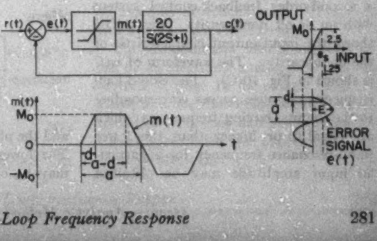


Fig. 13. Block diagram, nonlinearity, curves e(t) and m(t) of second-order feedback-control system with saturation nonlinearity

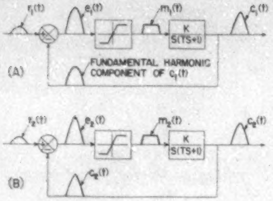


Fig. 14. Waveforms of various signals of nonlinear feedback-control system with saturation

A—When feedback signal is assumed to be fundamental-harmonic component of output
B—When feedback signal is exact output

not higher than the upper jump-resonance frequency ω_{f2}, the linear theory is used to find the attenuation-frequency and phase-frequency characteristics for the system under linear operation. The lower jump-resonance frequency ω_{f1} may be found by substituting equation 10 and the relation $a_1=\pi/\omega_{f1}$ into equation 11 in the following manner

$$\frac{T}{d}\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\times\tanh\frac{\pi}{2T\omega_{f1}}-1\right)+\frac{\pi}{2T\omega_{f1}}-\frac{d}{2T}=\frac{A}{KM_oT} \qquad (14)$$

where ω_{f1} is obtained in rad/sec (radians per second).

The upper jump-resonance frequency ω_{f2} may be found from equation 13 as follows:

$$\left|\frac{Aj\omega_{f2}(Tj\omega_{f2}+1)}{j\omega_{f2}(Tj\omega_{f2}+1)+M_oK/e_s}\right|=e_s \qquad (15)$$

When d is small compared with a, the half-period of the output, the approximation by setting $d=0$ may be suggested in order to simplify the numerical calculation unless a very high accuracy is needed. When d is set to be zero, equations 16 through 19 should be used instead of equations 6 through 8 and 14.

SECOND-ORDER FEEDBACK-CONTROL SYSTEM WITH ON-OFF NONLINEARITY

Fig. 10(A) shows the block diagram of a second-order feedback-control system with an on-off nonlinearity. Fig. 10(B) shows the input-output characteristic of the nonlinearity. The waveform of $m(t)$ is shown in Fig. 10(C). The closed-loop frequency-response curves corresponding to the inputs having frequencies which are equal to or higher than the lower jump-resonance frequency for a particular input amplitude may be obtained from the three equations which follow

$$\frac{C_p}{A}=\frac{KM_oT}{A}\ln\cosh\frac{\pi}{2T\omega} \qquad (16)$$

$$\phi=180°\left(1-\frac{T\omega}{\pi}\ln\cosh\frac{\pi}{2T\omega}\right)-\sin^{-1}\frac{KM_oT}{A}\left(\frac{\pi}{2T\omega}-\tanh\frac{\pi}{2T\omega}\right) \qquad (17)$$

where

$$0°\leqslant\sin^{-1}\left[\frac{KM_oT}{A}\left(\frac{\pi}{2T\omega}-\tanh\frac{\pi}{2T\omega}\right)\right]\leqslant 90° \qquad (18)$$

and the phase lag ϕ is obtained in degrees. The lower jump-resonance frequency ω_{f1} may be obtained from the following equation

$$\frac{\pi}{2T\omega_{f1}}-\tanh\frac{\pi}{2T\omega_{f1}}=\frac{A}{KM_oT} \qquad (19)$$

where ω_{f1} is obtained in rad/sec.

SECOND-ORDER FEEDBACK-CONTROL SYSTEM WITH ON-OFF AND DEAD-ZONE NONLINEARITY

The block diagram of a second-order feedback-control system with an on-off and dead-zone nonlinearity is shown in Fig. 11(A). The input-output characteristic and the output waveform $m(t)$ of the nonlinearity are shown in Figs. 11(B) and (C) respectively. When the maximum of the system output during the half-cycle $(0\leqslant\tau\leqslant a)$ at the steady state occurs during the time interval $b\leqslant\tau\leqslant a-b$, where b is the time corresponding to the dead-zone amplitude as shown in Fig. 11(C), the closed-loop frequency-response curves corresponding to the inputs having frequencies which are equal to or higher than the lower jump-responance frequency for a particular input amplitude may be obtained from the following equations

$$\frac{C_p}{A}=\frac{KM_oT}{A}\left(\ln\cosh\frac{\pi}{2T\omega}-\ln\cosh\frac{b}{T}\right) \qquad (20)$$

$$\phi=180°\left[1-\frac{T}{a}\left(\ln\cosh\frac{\pi}{2T\omega}-\ln\cosh\frac{b}{T}\right)\right]-\sin^{-1}\left[\frac{KM_oT}{A}\left(\frac{\pi}{2T\omega}-\frac{b}{T}-\cosh\frac{b}{T}\times\tanh\frac{\pi}{2T\omega}+\sinh\frac{b}{T}\right)\right] \qquad (21)$$

where

$$0°\leqslant\sin^{-1}\left[\frac{KM_oT}{A}\left(\frac{\pi}{2T\omega}-\frac{b}{T}-\cosh\frac{b}{T}\times\tanh\frac{\pi}{2T\omega}+\sinh\frac{b}{T}\right)\right]\leqslant 90° \qquad (22)$$

and the phase lag ϕ is obtained in degrees. The lower jump-resonance frequency ω_{f1} may be obtained from the following:

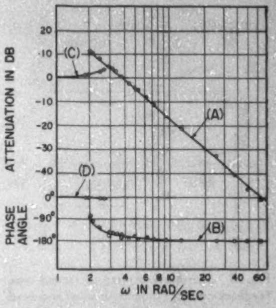


Fig. 15. Attenuation-frequency and phase-frequency characteristics of system shown in Fig. 13

$$\frac{\pi}{2T\omega_{f1}}-\cosh\frac{b}{T}\times\tanh\frac{\pi}{2T\omega_{f1}}-\frac{b}{T}+\sinh\frac{b}{T}=\frac{A}{KM_oT} \qquad (23)$$

SECOND-ORDER FEEDBACK-CONTROL SYSTEM WITH SATURATION AND DEAD-ZONE NONLINEARITY

The block diagram of a second-order feedback-control system with a saturation and dead-zone nonlinearity is shown in Fig. 12(A). Figs. 12(B) and (C) show the input-output characteristic and the output waveform of the nonlinearity. When the maximum of the system output during the half-cycle $(0\leqslant\tau\leqslant a)$ at the steady state occurs during the time interval $d\leqslant\tau\leqslant a-d$, shown in Fig. 12(C), the closed-loop frequency-response curves corresponding to the inputs having frequencies which are equal to or higher than the lower jump-resonance frequency for a particular input amplitude may be obtained from the following equations

$$\frac{C_p}{A}=\frac{KM_oT}{A}\left[\ln\frac{d-b}{T}+\ln\cosh\frac{\pi}{2T\omega}-\ln\left(\sinh\frac{d}{T}-\sinh\frac{b}{T}\right)\right] \qquad (24)$$

$$\phi=180°\left\{1-\frac{T\omega}{\pi}\left[\ln\frac{d-b}{T}+\ln\cosh\frac{\pi}{2T\omega}-\ln\left(\sinh\frac{d}{T}-\sinh\frac{b}{T}\right)\right]\right\}-\delta \qquad (25)$$

where

$$\delta=\sin^{-1}\left(\frac{KM_oT}{A}\left\{\frac{\pi}{2T\omega}-\frac{d+b}{2T}-\frac{T}{d-b}\times\left[\cosh\frac{b}{T}-\cosh\frac{d}{T}+\tanh\frac{\pi}{2T\omega}\times\left(\sinh\frac{d}{T}-\sinh\frac{b}{T}\right)\right]\right\}\right),\ 0°\leqslant\delta\leqslant 90° \qquad (26)$$

and the phase lag ϕ is obtained in degrees.

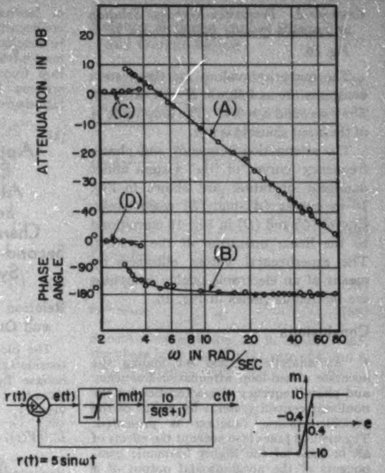


Fig. 16. Attenuation-frequency and phase-frequency characteristics of second-order feedback-control system with saturation nonlinearity

The lower jump-resonance frequency ω_{j1} may be obtained from the following:

$$\frac{\pi}{2T\omega_{j1}}-\frac{T}{d-b}\tanh\frac{\pi}{2T\omega_{j1}}\times\left(\sinh\frac{d}{T}-\sinh\frac{b}{T}\right)-\frac{d+b}{2T}-\frac{T}{d-b}\left(\cosh\frac{b}{T}-\cosh\frac{d}{T}\right)=\frac{A}{KM_oT} \quad (27)$$

where ω_{j1} is obtained in rad/sec.

Example 1. Frequency-Response Solutions of Feedback-Control System Shown in Fig. 13.

Approximate Solution. Assume that d is equal to zero. In this example $K=20$, $T=2$ sec, $M_o=2.5$, and $e_s=1.25$. The system input $r(t)$ is assumed to be a sinusoid $A \sin \omega t$, where $A=2$. Since the equations giving the analytic frequency-response solutions should be used for the inputs having frequencies equal to or higher than the lower jump-resonance frequency corresponding to the particular input amplitude, the lower jump-resonance frequency is found first. By substituting numerical values of A, K, T, and M_o into equation 19, the lower jump-resonance frequency ω_{j1} may be found from the following equation

$$\frac{0.785}{\omega_{j1}}-\tanh\frac{0.785}{\omega_{j1}}=0.02 \quad (28)$$

By using a mathematical table, it may be found that tanh χ is equal to 0.37995 when χ is equal to 0.40. Therefore, when

$$\frac{0.785}{\omega_{j1}}=0.40 \quad (29)$$

equation 28 is almost satisfied. From equation 29

$$\omega_{j1}=1.96 \text{ rad/sec} \quad (30)$$

When the input frequency is equal to or higher than 1.96 rad/sec, the attenuation-frequency and phase-frequency characteristics may be obtained from equations 16 through 18. The upper jump-resonance frequency may be found by substituting the numerical values of A, K, T, M_o, and e_s into equation 15. The upper jump-resonance frequency ω_{j2} thus found is 2.77 rad/sec. When the input frequency is equal to or less than 2.77 rad/sec, the attenuation-frequency and phase-frequency characteristics may be obtained from the linear theory.

Accurate Solution. Equations 6 through 8 should be used for finding accurate attenuation-frequency characteristics and phase-frequency characteristics. It should be noted here that when the analytic method presented in this paper is used for finding frequency-response characteristics, the feedback signal can be either sinusoidal or nonsinusoidal. Fig. 14(A) shows the waveforms of the signals $r_1(t)$, $e_1(t)$, $m_1(t)$, and $c_1(t)$ when the feedback signal is the fundamental-harmonic component of $c_1(t)$. Fig. 14(B) shows the waveforms of the signals $r_2(t)$, $e_2(t)$, $m_2(t)$, and $c_2(t)$ when the exact output $c_2(t)$ is fed back. The waveforms of $e_1(t)$ and $e_2(t)$ are different from each other, but the waveforms $m_1(t)$ and $m_2(t)$ are almost the same. Therefore, to simplify the numerical calculation, the feedback signal may be assumed to be sinusoidal. The amplitude of the feedback sinusoid may be assumed to be either the amplitude C_p of the system output or the amplitude of the fundamental-harmonic component of the system output.

Since equations 6 through 8 are functions of ω and d which is a function of a and E, the amplitude of the error signal; the values of E for various values of a or ω must be found first. Assume that the feedback signal is a sinusoid with an amplitude of C_p. Then the following equation may be obtained from Fig. 13

$$d=\frac{a}{\pi}\sin^{-1}\left(\frac{e_s}{E}\right) \quad (31)$$

where $a/\pi=1/\omega$

By simple analysis the following equation may also be obtained

$$E=\sqrt{A^2-2AC_p\cos\phi+C_p^2} \quad (32)$$

where ϕ is the output phase lag. In equation 32, C_p and ϕ may be obtained from equations 16 through 18. From equations 31 and 32, the values of d for various values of a or ω may be found. Then the accurate attenuation-frequency and phase-frequency characteristics for the system under nonlinear operation may be found from equations 6 through 8 by using the values of d thus obtained. The accurate lower jump-resonance frequency ω_{j1} may be found from equation 14. The value of ω_{j1} obtained from equation 14 is 1.99 rad/sec. The attenuation-frequency and phase-frequency curves obtained by use of equations 6 through 8 are shown in Fig. 15 as curves (*A*) and (*B*) respectively. The attenuation-frequency and phase-frequency curves corresponding to the linear operation of the system are shown in Fig. 15 as curves (*C*) and (*D*) respectively. The results obtained by the analytic method were compared with those obtained by an electronic-analog-computer study. The values of the lower jump-resonance frequency and the upper jump-resonance frequency obtained by the analytic method are 1.99 rad/sec and 2.77 rad/sec respectively, and those obtained by experiments are 2.01 rad/sec and 2.75 rad/sec respectively.

The experimental results obtained by use of an electronic-analog-computer are plotted in Fig. 15 as circles. Fig. 15 shows that the theoretical closed-loop attenuation-frequency and phase-frequency curves agree with the experimental results satisfactorily.

Example 2. Frequency-Response Solution of Feedback-Control System Shown in Fig. 16.

The numerical values for the system constants are as follows: $K=10$, $M_o=10$, $T=1$ sec, and $e_s=0.4$. The amplitude A of the input sinusoid is five.

The attenuation-frequency and phase-frequency curves of the system under nonlinear operation are shown in Fig. 16 as curves (A) and (B) respectively. Curves (C) and (D) in Fig. 16 correspond to the linear operation of the system. The experimental values obtained by means of an electronic analog computer are plotted as circles in Fig. 16.

Conclusions

1. An analytic method for finding the accurate closed-loop attenuation-frequency and phase-frequency characteristics for the nonlinear control systems by means of the periodic-response function is presented. The method takes into account the effects of all or most of the higher harmonic components of the nonsinusoidal output of a nonlinear element for finding the system output.

2. The analytic closed-loop frequency-response solutions for fundamental oscillations are shown for the following systems:
   a. The second-order feedback-control system with the saturation nonlinearity shown in Fig. 3.
   b. The second-order feedback-control system with the on-off nonlinearity shown in Fig. 10.
   c. The second-order feedback-control system with the on-off and dead-zone nonlinearity shown in Fig. 11.
   d. The second-order feedback-control system with the saturation and dead-zone nonlinearity shown in Fig. 12.

The closed-loop attenuation-frequency and phase-frequency characteristics for the foregoing systems for fundamental oscillations may be obtained by merely substituting system constants and input amplitudes into corresponding equations shown in the paper. Since the equations developed in this paper giving amplitudes and phase lags of the system output as functions of the output frequency contain the system constants as parameters, the contribution of these constants to the attenuation-frequency and phase-frequency characteristics is evident. Therefore, these equations may be used for synthesizing corresponding nonlinear feedback-control systems.

3. Methods for finding the jump resonance frequencies in terms of the system constants and the input amplitude are presented. The jump resonance frequencies obtained by these methods agree quite satisfactorily with the experimental values obtained by means of an electronic analog computer.

4. The closed-loop frequency-response characteristics of nonlinear feedback-control systems obtained by the present method agree quite favorably with the experimental results obtained by an electronic-analog-computer study. Since the equations for finding the closed-loop frequency response of feedback-control systems with on-off nonlinearities are not complicated, it may be concluded that this method may be used most advantageously for finding the closed-loop frequency-response characteristics of systems having on-off nonlinearities or high-degree saturation nonlinearities.

Appendix. Derivation of Equations for Finding Attenuation-Frequency and Phase-Frequency Characteristics of Nonlinear Second-Order Feedback-Control System with Saturation

Relation Between Output Velocity and Output Period

The block diagram of the system under consideration is shown in Fig. 3(A). The inverse Laplace transform of equation 4 gives the time solution of the system output velocity.

$$L^{-1}[V(s)]=v(t)$$

$$=\frac{KM_o}{d}\left\{t-\sum_{n=0}^{\infty}(-1)^n(t-d-an)\times u(t-d-an)-\sum_{n=0}^{\infty}(-1)^n(t+d-a-an)\times u(t+d-a-an)-T+T\sum_{n=0}^{\infty}\times (-1)^n\times u(t-d-an)+T\sum_{n=0}^{\infty}(-1)^n\times u(t+d-a-an)+T\left[\epsilon^{-t/T}-\sum_{n=0}^{\infty}(-1)^n\epsilon^{-\frac{1}{T}(t-d-an)}\times u(t-d-an)-\sum_{n=0}^{\infty}(-1)^n\epsilon^{-\frac{1}{T}(t+d-a-an)}\times u(t+d-a-an)\right]\right\}+v(0)\epsilon^{-t/T} \quad (33)$$

where

$$\left.\begin{array}{ll} u(t)=0 & \text{when } t<0 \\ u(t)=1 & \text{when } t>0 \end{array}\right\} \quad (34)$$

Equation 33 contains the initial velocity $v(0)$ term. To obtain $v(t)$ in a closed form and also to eliminate the $v(0)$ term in equation 33, let

$$\lambda a\leqslant t\leqslant(\lambda+1)a \quad (35)$$

where λ is a sufficiently large number. Then $\epsilon^{-t/T}$ becomes negligibly small, with the result that the initial condition term in equation 33 may be eliminated.

$$v(0)\epsilon^{-t/T}\rightarrow 0 \quad (36)$$

Arbitrarily choosing λ as an odd number, the value of $m(t)$ for $\lambda a<t<(\lambda+1)a$ becomes negative.

$$m[\lambda a<t<(\lambda+1)a]<0 \quad (37)$$

Fig. 17 shows curve $m(t)$. Since curve $m(t)$ consists of straight-line segments as shown in Fig. 17, the time interval $\lambda a\leqslant t\leqslant(\lambda+1)a$ may be divided into the following three intervals corresponding to the three straight-line segments.

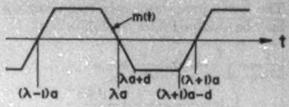


Fig. 17. Curve m(t)

$$\begin{array}{l} \lambda a\leqslant t\leqslant\lambda a+d \\ \lambda a+d\leqslant t\leqslant(\lambda+1)a-d \\ (\lambda+1)a-d\leqslant t\leqslant(\lambda+1)a \end{array} \quad (38)$$

To simplify the derivation of equations, let

$$t-\lambda a\equiv\tau \quad (39)$$

and substitute τ in equation 38, thus the three time intervals may be written as

$$\left.\begin{array}{l} 0\leqslant\tau\leqslant d \\ d\leqslant\tau\leqslant a-d \\ a-d\leqslant\tau\leqslant a \end{array}\right\} \quad (40)$$

The output velocity $v(\tau)$ for each time interval may be obtained as a function of τ, a, the half-period of the output, and d, as shown in the following.

For Time Interval $0\leqslant\tau\leqslant d$

By applying equation 34 to equation 33, the upper limits of the summation signs in equation 33 become $\lambda-1$ instead of ∞ for this time interval. Then equation 33 may be simplified as follows

$$v(0\leqslant\tau\leqslant d)=v_1(\tau)$$

$$=\frac{KM_o}{d}\left\{(\tau+\lambda a)-\sum_{n=0}^{\lambda-1}(-1)^n(\tau+\lambda a-d-an)-\sum_{n=0}^{\lambda-1}(-1)^n(\tau+\lambda a+d-a-an)-T+2T\sum_{n=0}^{\lambda-1}(-1)^n+T\left[-\sum_{n=0}^{\lambda-1}(-1)^n\epsilon^{-\frac{1}{T}(\tau+\lambda a-d-an)}-\sum_{n=0}^{\lambda-1}(-1)^n\epsilon^{-\frac{1}{T}(\tau+\lambda a+d-a-an)}\right]\right\}$$

$$=\frac{KM_o}{d}\left[-\tau+T-T\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{a}{2T}\right)\epsilon^{-\tau/T}\right] \quad (41)$$

For Time Interval $d\leqslant\tau\leqslant a-d$

$$v(d\leqslant\tau\leqslant a-d)=v_2(\tau)$$

$$=KM_o\left(-1+\frac{T}{d}\epsilon^{\frac{a}{2T}-\frac{\tau}{T}}\times\frac{\sinh\frac{d}{T}}{\cosh\frac{a}{2T}}\right) \quad (42)$$

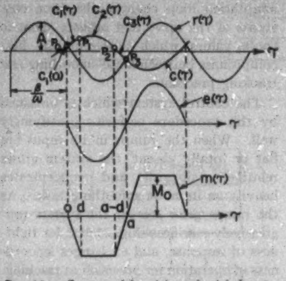


Fig. 18. Curves $c_1(\tau)$, $c_2(\tau)$, and $c_3(\tau)$ showing continuous and symmetric nature of $c(\tau)$

For Time Interval $a-d\leq\tau\leq a$

$$v(a-d\leq\tau\leq a)=v_3(\tau)$$
$$=\frac{KM_0}{d}\left[(\tau-a)-T+T\times\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{a}{2T}\right)\epsilon^{-\frac{\tau-a}{T}}\right] \quad (43)$$

Relation Between Output Position and Output Period

To find the expression for the output amplitude as a function of the output frequency, the output position is found as a function of τ, a, and d.

Since

$$c_i(\tau)=\int v_i(\tau)d\tau,\ i=1, 2, 3 \quad (44)$$

$c_1(\tau)$, $c_2(\tau)$, and $c_3(\tau)$ may be found by integrating equations 41 through 43 with respect to τ. Because of the continuous and symmetric nature of $c(\tau)$ as shown in Fig. 18, the following equations are obtained.

At point P_1 in Fig. 18

$$c_1(\tau\to d)=c_2(\tau\to d) \quad (45)$$

At point P_2

$$c_2(\tau\to a-d)=c_3(\tau\to a-d) \quad (46)$$

At points P_0 and P_3, $c_1(\tau)$ and $c_3(\tau)$ have the same magnitudes with the opposite signs. Therefore

$$c_1(\tau\to 0)=-c_3(\tau\to a) \quad (47)$$

Using equations 45 through 47, the system output may be found as follows

$$c(0\leq\tau\leq d)=c_1(\tau)=\frac{KM_0T}{d}\left[\tau-T+T\times\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{a}{2T}\right)\epsilon^{-\tau/T}\right]+\frac{KM_0}{2}(a-d)-\frac{KM_0\tau^2}{2d} \quad (48)$$

$$c(d\leq\tau\leq a-d)=c_2(\tau)=KM_0T\times\left(-\frac{\tau}{T}-\frac{T}{d}\epsilon^{\frac{a}{2T}-\frac{\tau}{T}}\times\frac{\sinh\frac{d}{T}}{\cosh\frac{a}{2T}}\right)+KM_0\left(\frac{a}{2}+T\right) \quad (49)$$

$$c(a-d\leq\tau\leq a)=c_3(\tau)=\frac{KM_0T}{d}\left[T-T\times\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{a}{2T}\right)\epsilon^{\frac{a-\tau}{T}}-(\tau-a)+\frac{(\tau-a)^2}{2T}\right]-\frac{KM_0}{2}(a-d) \quad (50)$$

Substituting $\tau=0$ in equation 48, equation 10 may be obtained. Equations 48 through 50 show $c_1(\tau)$, $c_2(\tau)$, and $c_3(\tau)$ as functions of τ and a, the half-period of the output, and d. The amplitude of the output $c(\tau)$ may be found as the maximum of $c(0\leq\tau\leq a)$ during the time interval $0\leq\tau\leq a$ as shown in the following.

Relation Between Output Amplitude and Output Period

The maximum of $c(\tau)$ occurs when $v(\tau)$ becomes zero. If d/a is small and the velocity of the output becomes zero at some instant between the time interval d and $a-d$, the maximum of $c(\tau)$ occurs during the time interval $d\leq\tau\leq a-d$. Since τ_{max} was defined before as the time when the output position $c(\tau)$ becomes maximum, τ_{max} may be obtained by letting $v_2(\tau)=0$.

$$\tau_{max}=T\left(\ln\sinh\frac{d}{T}-\ln\frac{d}{T}-\ln\cosh\frac{a}{2T}+\frac{a}{2T}\right) \quad (51)$$

Let C_p be the output amplitude. Substituting equation 51 in equation 49, C_p or $c_2(\tau_{max})$ may be written as follows:

$$C_p=c_2(\tau_{max})=KM_0T\times\left(\ln\cosh\frac{a}{2T}-\ln\sinh\frac{d}{T}+\ln\frac{d}{T}\right) \quad (52)$$

Equation 52 shows the output amplitude C_p as a function of a and d.

Relation Between Output Phase Lag and Output Period

By substituting $\tau=0$ in equation 48, equating $c(0)$ to $A\sin\beta$, and letting $\delta=180°-\beta$, the following equation may be obtained:

$$\sin\delta=\frac{KM_0T}{Ad}\left[-T+T\times\left(\cosh\frac{d}{T}-\sinh\frac{d}{T}\tanh\frac{a}{2T}\right)+\frac{(a-d)d}{2T}\right] \quad (53)$$

From Fig. 4,

$$\theta_1=90°-\delta \quad (54)$$

$$\theta_2=\tau_{max}\omega \quad (55)$$

Substituting equation 51 in equation 55, the phase lag ϕ which is the sum of θ_1 and θ_2 may be obtained as follows:

$$\phi=180°\left[1-\frac{T}{a}\left(\ln\cosh\frac{a}{2T}+\ln\frac{d}{T}-\ln\sinh\frac{d}{T}\right)\right]-\delta \quad (56)$$

Substituting $a=\pi/\omega$ in equations 53 and 56, equations 7 and 8 may be obtained.

References

1. NONLINEAR CONTROL SYSTEMS, N. Minorsky. "Automatic and Manual Control." *Proceedings*, Cranfield Conference 1951, Butterworths Scientific Publications, London, England, 1952, pp. 309–18.

2. SOME SATURATION PHENOMENA IN SERVOMECHANISMS WITH EMPHASIS ON THE TACHOMETER STABILIZED SYSTEM, E. Levinson. *AIEE Transactions*, vol. 72, pt. II, Mar. 1953, pp. 1–9.

3. A GENERALIZED METHOD FOR DETERMINING THE CLOSED-LOOP FREQUENCY RESPONSE OF NONLINEAR SYSTEMS, Luther T. Prince, Jr. *Ibid.*, vol. 73, pt. II, Sept. 1954, pp. 217–24.

4. AUTOMATIC FEEDBACK CONTROL SYSTEM SYNTHESIS (book), J. G. Truxal. McGraw-Hill Book Company, Inc., New York, N. Y., 1955.

5. A FREQUENCY RESPONSE METHOD FOR ANALYZING AND SYNTHESIZING CONTACTOR SERVOMECHANISMS, Ralph J. Kochenburger. *Ibid.*, vol. 69, pt. I, 1950, pp. 270–84.

6. SINUSOIDAL ANALYSIS OF FEEDBACK-CONTROL SYSTEMS CONTAINING NONLINEAR ELEMENTS, E. Calvin Johnson. *Ibid.*, vol. 71, pt. II, July 1952, pp. 169–81.

7. SUB- AND SUPER-HARMONIC OSCILLATIONS OF NONLINEAR CONTROL SYSTEMS, K. Ogata. *Ph.D. Thesis*, University of California, Berkeley, Calif., June 1956.

8. MODERN OPERATIONAL MATHEMATICS IN ENGINEERING (book), R. V. Churchill. McGraw-Hill Book Company, Inc., 1944, p. 54.

Adaptive Servomechanisms

R. F. DRENICK
NONMEMBER AIEE

R. A. SHAHBENDER
NONMEMBER AIEE

Synopsis: A generalized approach is developed to aid in the synthesis of servomechanisms exhibiting adaptive behavior. The design methods described in this paper permit the synthesis of adaptive systems to follow polynomial signals up to any predetermined degree in the presence of Gaussian noise of known autocorrelation function. The basis of the method is the requirement that the servo loop adjust its parameters so as to minimize the steady-state root-mean-square error resulting from a polynomial input in the presence of noise. This condition specifies the functional relation between the servo parameters and the polynomial input. The physical realization of these functional relationships is treated in the paper and leads to the adaptation loops necessary to produce the adaptive behavior. As an example of the synthesis procedure, the design of a second-order adaptive system is presented. Analog computer results demonstrating that the adaptive system is superior in performance to linear systems of the same order are also included.

THE PURPOSE of this paper is to present an approach to the logical design of a class of servomechanisms which we have here designated as "adaptive." This term is borrowed from biology where it describes the ability of an organism to adjust itself to its environment. It is here used in the same sense. Specifically it refers to control systems which, roughly speaking, monitor their own performance and adjust some of their parameters in the direction of better performance.

It may be worth pointing out that such adjustments, whether in an organism or in a control system, need not be brought about by the self-monitoring principle suggested in the foregoing. It is often possible to monitor the environment directly and to adjust the control system accordingly. For example in the case of an aircraft auto-pilot, it is possible to determine the altitude of flight and to continuously adjust the auto-pilot for optimum performance. However, there exist situations in which the independent sensing of the environment is impractical or impossible. It is such cases that are dealt with in this paper.

The problem is more specifically the following. A control system is to be designed for a load with inertia and viscous damping, supplemented with a suitable equalizing network so that it can properly follow a signal embedded in noise. The signal, in Part II of the paper, is assumed to be a ramp function. The changing environment which can lead to the requirement for adaptation, is introduced by the fact that the slope of the ramp changes from time to time.

The manner in which the requirement for adaptation arises can be seen by the following argument. In choosing a suitable equalizing network, the designer is faced with a dilemma. He knows that it is necessary to include proportional plus integral action in the equalizing network in order to avoid steady-state tracking errors. In the case where the command signal is not contaminated with noise, the designer would undoubtedly choose such a mode of control. The parameters of the equalizing network are then chosen for acceptable transient performance.

When the incoming signal is contaminated with noise, the problem is more complicated. An equalizing network giving proportional control is more effective in smoothing the noise, than a network giving proportional plus integral action. However, the servo system with pure proportional control is not capable of properly tracking a ramp input. It is afflicted with a steady state tracking error which offsets its superior noise smoothing ability.

A compromise is clearly desirable. The adaptive servomechanisms proposed in this paper strive, in a way, to attain this compromise. The simple proportional control is retained for the equalizing network. However, unlike a conventional servo, an additional loop, called the "adaptation loop" is added to the system. The purpose of this loop is to monitor the performance of the servo and to detect tracking errors in its output which are attributable to an inadequately followed ramp input, rather than to mere noise. When such an error is detected the adaptation loop changes the time constants of the servo and adjusts them to those values which represent the best compromise between noise-smoothing and tracking precision.

The control system which is obtained by this procedure performs gratifyingly well. When the ramps in its input are flat or totally absent, the system grows relatively sluggish and concentrates heavily on its noise smoothing task. As the ramps grow steeper the system progressively sacrifices sluggishness for tightness of response, and exchanges smoothness of operation for precision in tracking.

The design procedure developed in this paper pertains to a class of synthesis problems which is similar to that treated by Phillips.[1] However, the assumptions made in this paper and the results obtained are different. It is hoped, more particularly, that the procedure developed will be more easily carried out in practice and lead to mechanisms whose all-round performance is less controversial. This might make up for some of the conceptual shortcomings of the theory.

The paper is organized as follows: Part II contains the exposition in detail for the synthesis of the adaptive servo which has been used previously as the example for discussion. This is the servo with simple proportional control, a second-order servo, supplemented with an adaptation loop. The idea of the proper compromise with noise smoothing and precision tracking is first taken up. This is followed by a method for the design of the adaptation loop.

The case of the second-order servo is virtually the simplest conceivable case to which this procedure can be applied; yet it contains most of the relevant features of the general case. To facilitate the transition to the latter, an intermediate case is treated next, in Part III. This is the third-order servo, that is to say, a servomechanism in which the equalizing network incorporates proportional and integral control, both adjustable by an adaptation loop. The common features of the two examples point directly to the general case which is discussed briefly (Part IV).

In order to illustrate the usefulness of the procedure, a numerical example is discussed. Part V contains the analytical evaluation, as far as possible, of a second-order adaptive servo, along with two borderline competitors: conventional linear servos with simple proportional control, one synthesized with emphasis on noise-smoothing, the other on precision tracking. Part VI contains a discussion of some analog computer runs on the same

---

Paper 57-388, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted July 5, 1956; made available for printing April 24, 1957.

R. F. DRENICK is with Bell Telephone Laboratories, Inc., New York, N. Y., and R. A. SHAHBENDER is with the Radio Corporation of America, Camden, N. J.

The authors wish to acknowledge the interested assistance of the staff of Electronics Associates Computation Center, Princeton, N. J., in setting up the computer simulation and in obtaining the experimental results.

This work was supported in part by the U. S. Air Force under Contract *AF 33(600)28007*.

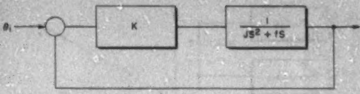


Fig. 1 (left). Proportional control system

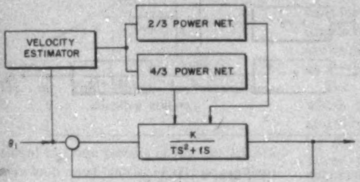


Fig. 2 (left). Second-order adaptive servo system

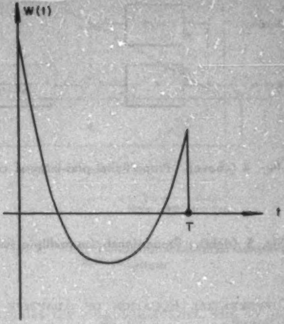


Fig. 3 (right). Optimum weighting function for velocity estimator

examples. Both discussions show that the adaptive servo combines the advantages of both competitors to a gratifying degree.

II. Second-Order Systems

Preliminary Remarks

In this section an outline will be given of the approach to the design of an adaptation loop for a positional servo with proportional control (Fig. 1), that is, a servo whose linear version is described by the second-order transfer function

$$\frac{\theta_o(s)}{\theta_i(s)}=\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \tag{1}$$

$$\omega_n^2=K/J$$

$$\zeta=f/2\sqrt{KJ}$$

where

$\theta_i(t)$ = input to servo system
$\theta_o(t)$ = output of servo system
ω_n = angular frequency of undamped oscillation
ζ = damping ratio
J = motor and load inertia
f = viscous damping coefficient
K = motor amplifier gain (proportional gain or spring constant)
s = Laplace variable

The input to the servo, that is, to the linear as well as to the adaptive versions, consists of a ramp signal

$$S(t)=\begin{cases}A_r t & (t\geq 0)\\ 0 & (t<0)\end{cases} \tag{2}$$

and Gaussian noise of known autocorrelation function $R_N(\tau)$. In fact, since the design procedure to be outlined in the following is feasible regardless of the form of the autocorrelation function, it will be sufficient to assume for it its simplest form

$$R_N(\tau)=(1/2)\sigma_N^2\delta(\tau) \tag{3}$$

corresponding to white noise.

The performance of the servos, linear as well as adaptive, will be judged by the square error in the steady state. This error has two components: the mean-square error due to the noise (ϵ_N^2) and the square of the servo lag in following the signal(ϵ_s^2). The total square error (ϵ_T^2) is accordingly

$$\epsilon_T^2=\epsilon_s^2+\overline{\epsilon_N^2} \tag{4}$$

and will serve in what follows as performance and optimization criterion.[2]

Parameter Variation of Second-Order Systems

For the linear second-order proportional control system defined in the foregoing, the total error in equation 4 in following a ramp input is

$$\epsilon_T^2=\frac{4\zeta^2}{\omega_n^2}A_r^2+\frac{\omega_n}{8\zeta}\sigma_N^2$$

For fixed ζ, the minimum value of ϵ_T^2 is attained when:

$$\omega_n=\frac{4\zeta}{(\sigma_N^2)^{1/3}}A_r^{2/3}$$

and if, in addition, the inertia is fixed, this requires

$$K\alpha A_r^{4/3},\ f\alpha A_r^{2/3} \tag{5}$$

Therefore, to have optimum operation, the slope of the ramp signal would have to be determined, a task which is impossible, strictly speaking, because the noise overlies the signal. Instead of an exact determination, however, one can resort to a statistical estimation procedure. This will be relegated to the adaptation loop.

Optimum Adaptation Loop for Second-Order Systems

The adaptation loop for the second-order system consists of a first-derivative (or velocity) estimator, followed by networks that give the required powers (equation 5) of the estimator output as shown in Fig. 2. Thus the adaptation loop design problem reduces to the choice of a suitable velocity estimator. The weighting function $W(t)$ of the optimum estimator, which may be determined by the method of Zadeh and Ragazzini,[3] leads, as shown in Appendix I, to the following result:

$$W(t)=\frac{36}{T^2}-\frac{192}{T^3}t+\frac{180}{T^4}t^2 \quad 0\leq t\leq T$$

$$=0 \qquad \text{elsewhere.} \tag{6}$$

Fig. 3 is a plot of the foregoing function. Since $W(t)$ is discontinuous at $t=T$, it is not realizable in terms of lumped elements; and it becomes necessary, as will be discussed later, to determine a realizable approximation to this weighting function.

The behavior of the optimum estimator in equation 6 is evidently governed by the choice of T, which can be effected by the following line of reasoning. Since the output of the estimator attains its steady-state value after an interval of T seconds from the application of a ramp input, it is desirable to make T as short as possible. On the other hand, the mean-square noise σ_E^2 at the output of the estimator network is

$$\sigma_E^2=\frac{96}{T^3}\sigma_N^2$$

which implies as large a T as possible.

From these considerations it is possible to fix a suitable compromise value for T. As shown in the appendix, such a value is

$$T=\left[\frac{96\alpha_1^2\sigma_N^2}{A_1^2}\right]^{1/3}$$

where

α_1 = minimum acceptable signal-to-noise ratio for the estimator
A_1 = smallest value of input velocity to be detected

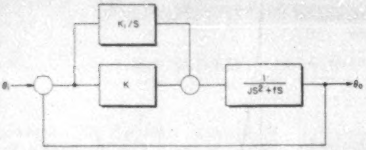


Fig. 4 (above). Proportional-plus-integral control system

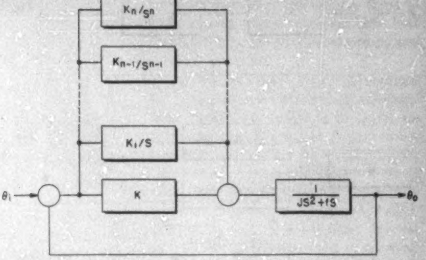


Fig. 5 (right). Proportional-plus-multiple integral control system

Differential Equation of Adaptive Servo

The adaptive servo just synthesized obeys a differential equation of the form

$$\ddot{\theta}_o+C_1\theta_E{}^{2/3}\dot{\theta}_o+C_2\theta_E{}^{4/3}\theta_o=C_3\theta_E{}^{4/3}\theta_i$$

where θ_E has been written for the output of the estimator network

$$\theta_E(t)=\int_{-\infty}^{t}W(\tau)\theta_i(t-\tau)d\tau$$

This is an interesting type of equation: It is linear in the sense that the dependent variable θ_o and its derivative enter linearly only; but it is nonlinear in that it violates the super-position principle.

Its stability is now an only partially answered question. It can be proved stable under a very large class of inputs, including step and ramp functions; but it is not known, for instance, whether stability is assured under sinusoidal inputs regardless of frequency.

Physically Realizable Adaptation Loop for Second-Order System

The weighting function of equation 6 of the optimum velocity estimator is not realizable in terms of lumped elements. To determine a weighting function $W_a(t)$ that approximates the optimum function and that is realizable in terms of lumped elements, apply Foster's t-method. Let

$$W_a(t)=\sum_{k=1}^{3}B_ke^{-\alpha_kt}$$

where the B_k's and α_k's are constants to be determined so that $W_a(t)$ approximates $W(t)$. The following set of conditions may be used to determine these constants.

1. $W_a(t)$ and $W(t)$ have equal moments.

$$\int_0^\infty t^iW_a(t)dt=\int_0^T t^iW(t)dt \quad i=0,\ 1,\ \text{and}\ 2$$

2. $W_a(t)$ has a tail area that is 1% of the total area:

$$\int_T^\infty W_a(t)dt=0.01\int_0^\infty W_a(t)dt$$

3. $W_a(0)=W(0)$

4. $W_a(t)$ has minimum mean-square noise output when the input consists of the noise signal $N(t)$.

It may be shown from the foregoing conditions that it is possible to determine the values of the constants associated with the weighting function and thus realize the estimator in terms of lumped elements.

An example of this procedure is given in Section V.

III. Third-Order Systems

Preliminary Remarks

The preceding discussion concerning the proportional control servo can readily be extended to the next case in order of complication: the positional servo with proportional-plus-integral control (Fig. 4). This is described by a third-order transfer function of the form:

$$\frac{\theta_o(s)}{\theta_i(s)}=\frac{(\omega_n{}^2+2\zeta\omega_n\gamma)s+\gamma\omega_n{}^2}{(s^2+2\zeta\omega_ns+\omega_n{}^2)(s+\gamma)}$$

$$2\zeta\omega_n+\gamma=f/J$$

$$2\zeta\omega_n\gamma+\omega_n{}^2=K/J$$

$$\gamma\omega_n{}^2=K_1/J$$

where

K = proportional gain
K_1 = integrator gain
γ = real pole of the transfer function

The input is now assumed to be a signal of the form

$$S(t)=\begin{cases}(1/2)A_qt^2 & (t\geq 0)\\ 0 & (t<0)\end{cases} \qquad (7)$$

contaminated with white Gaussian noise, as described above. The performance criterion is again the total steady state, squared error $\epsilon_T{}^2$.

Parameter Variation

For the third-order system, the total error in following the quadratic input of equation 7 is given by the expression:

$$\epsilon_T{}^2=\left[\frac{2\zeta\omega_n+\gamma}{\omega_n{}^2\gamma}\right]^2A_q{}^2+\frac{\omega_n}{8\zeta}\left[1+\frac{4\zeta\gamma(\omega_n+\zeta\gamma)}{\omega_n{}^2+2\zeta\omega_n\gamma+\gamma^2}\right]\sigma_N{}^2$$

The above expression includes three parameters (ζ, ω_n, and γ) that may, in principle, be varied to produce adaptive behavior. The step response of the third-order system consists of an exponential term plus a damped sinusoidal term. The damping of the sinusodial term is governed by ζ, which, having been chosen for good transient recovery, may be assumed fixed.

The parameter γ specifies the exponential term in the step response of the system. As shown in the literature,[1] the choice of

$$\gamma=2\zeta\omega_n$$

leads to a system with good performance characteristics. For that value of γ we have:

$$\epsilon_T{}^2=\frac{4}{\omega_n{}^4}A_q{}^2+\frac{\omega_n}{8\zeta}\sigma_N{}^2+\frac{1+2\zeta^2}{1+8\zeta^2}\zeta\omega_n\sigma_N{}^2$$

The minimum value of $\epsilon_T{}^2$ is then obtained when

$$\omega_n=2\left[\frac{4\zeta(1+8\zeta^2)}{(1+\zeta+8\zeta^2+2\zeta^3)\sigma_N{}^2}\right]^{1/5}A_q{}^{2/5}$$

Since J is assumed constant, and $\sigma_N{}^2$ is constant, the above expression may be written as:

$$\omega_n=CA_q{}^{2/5}$$

where C is a constant. Further, we have:

$$\gamma=2\zeta CA_q{}^{2/5}$$

From the relations between J, f, K, K_1 and ζ, ω_n, γ as given previously—again assuming unalterable inertia—one can determine the dependence on the acceleration A_q of the proportional gain K, the integrator gain K_1, and the viscous damping f of the servo system:

$$f\alpha A_q{}^{2/5},\qquad K\alpha A_q{}^{4/5},\qquad K_1\alpha A_q{}^{6/5}$$

Since the servo parameters are propor-

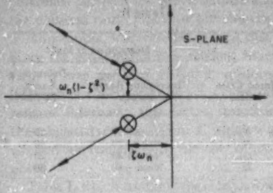


Fig. 6. Pole motion for second-order adaptive system

tional to these powers of A_2, the adaptation loop must consist of: (1) an acceleration estimator, and (2) networks to produce powers of the estimator output.

Optimum Adaptation Loop

The adaptation loop for the third-order system consists of an acceleration estimator in tandem with networks that give the required powers of the estimator output. The ideal weighting function of the estimator is given by:[3]

$$W(t)=\frac{480}{T^3}-\frac{5{,}400}{T^4}\,t+\frac{12{,}960}{T^5}\,t^2-\frac{8{,}400}{T^6}\,t^3 \qquad 0\le t\le T$$

$$=0 \quad \text{elsewhere}$$

which must, in practice, be again approximated by a physically realizable function.

The mean-square noise output of the estimator is given by:

$$\sigma_E^2=\frac{12.96\times10^3}{T^5}\,\sigma_N^2$$

As in the case of the second-order system, the estimator must have a minimum signal-to-noise ratio α_2 for a smallest detectable acceleration magnitude A_2. The estimator time constant is then given by:

$$T=\left[\frac{12.96(10^3\alpha_2^2\sigma_N^2}{A_2^2}\right]^{1/5}$$

IV. Generalization to Higher-Order Servomechanisms

The generalization of the notions exhibited above leads to a positional servo with proportional-plus-multiple-integral control (Fig. 5) with the transfer function

$$\frac{\theta_i(s)}{\theta_o(s)}=\frac{Ks^n+K_1s^{n-1}+\ldots+K_n}{Js^{n+2}+fs^{n+1}+Ks^n+K_1s^{n-1}+\ldots+K_n}$$

where

K = proportional gain
K_i = gain of ith integrator ($i=1, 2, \ldots n$)

An nth order servo system, as defined above, will exhibit no steady-state errors, other than noise errors, for polynomial inputs of degree less than, or equal to, $(n-2)$; constant steady-state errors for polynomial inputs of degree $(n-1)$; and steady-state errors that increase with time for inputs of degree n or higher. For inputs of degree n the steady-state error grows as t, and for inputs of degree $(n+1)$ as t^2— and so on.

The effect of the adaptation loops which were proposed above for second- and third-order systems can be illustrated by the motion of the poles, which is induced in the transfer functions of the linear system by the adaptive behavior. This idea can then be carried over to the design of higher-order systems.

A second-order servo system is characterized by a closed-loop transfer function having a single pair of complex conjugate poles in the s-plane. The adaptation loop operates so as to increase ω_n from its unadapted value in the presence of ramp inputs but maintaining ζ constant. Thus the adaptation loop operates so as to move the poles into the interior of the left-half of the complex plane along the two straight lines shown in Fig. 6. Similarly, in the case of a third-order system the closed-loop transfer function is characterized by a single pair of complex conjugate poles and a real pole. The adaptation loop operates so as to move the poles into the interior of the left half of the complex plane along the paths shown in Fig. 7.

A fourth-order system may be treated in the same manner—except for two cases that must be considered.

1. The closed-loop transfer function has two pairs of complex conjugate poles with corresponding damping ratios ζ_1, and ζ_2 and angular frequencies ω_1 and ω_2. The adaptation loop operates so as to move the poles in a manner similar to that shown in Fig. 6.
2. The closed-loop transfer function has a single pair of complex conjugate poles and two real poles, γ_1 and γ_2. For this case:

$$\gamma_1=C_1\omega_n$$

$$\gamma_2=C_2\omega_n$$

C_1 and C_2 are constants. The adaptation loop operates so as to move the poles in a manner similar to that shown in Fig. 7.

This can be generalized to systems of any order. However, it is of purely academic interest to consider, of the type under consideration, any systems higher than the fourth order.

For a second-order system the estimator consists of a network that determines the first derivative of the input. For a third-order system the estimator determines the second derivative, and so on. The procedure employed in determining the time constant associated with the adaptation loop for the second- and third-order systems may obviously be extended to systems of higher order.

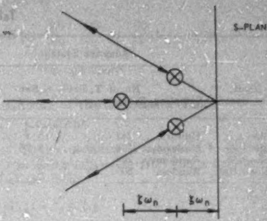


Fig. 7. Pole motion for third-order adaptive system

V. Design of a Specific Second-Order Adaptive System

For the second-order adaptive system the time constant associated with the velocity estimator is given by:

$$T=\left[\frac{96\alpha_1^2\sigma_N^2}{A_1^2}\right]^{1/3}$$

The relation between ω_n and A_r is given by:

$$\omega_n=\frac{4\zeta}{(\sigma_N^2)^{1/3}}A_r^{2/3}$$

Assuming the following values for the quantities given in the above expressions:

$\sigma_N^2=0.636$ square foot per second
$A_1=20$ feet per second
$\alpha_1=0.5$
$\zeta=0.5$

we obtain:

$T=0.534$ second
$\omega_n=2.33\,A_r^{2/3}$

The weighting function of the optimum estimator is completely determined by the value of T given above. The approximate weighting function may now be determined from the conditions given in Section II. One finds:

$$W_a(t)=250e^{-t/.089}-158e^{-t/.178}+15.8e^{-t/.356}$$

Let $\theta_E(t)$ denote the output of the estimator for an input $\theta_i(t)$. The adaptive-system damping and proportional gain are given by:

$$f/J=2.33\theta_E^{2/3}$$

$$K/J=5.43\theta_E^{4/3}$$

where J is assumed constant.

Table I

Input, Feet/Second	Adaptive System					Linear Tracker				
	Sampling	No. of Samples	T, Seconds	τ, Seconds	RMS Error, Feet	Sampling	No. of Samples	T, Seconds	τ, Seconds	RMS Error, Feet
Noise	Uniform	243		15.1	1.38	Uniform	200		5	2.65
Noise+20	Uniform	122	4.2	3.75	2.56	Uniform	121	4.2	3.75	2.69
Noise+50	Random	200	5	5	2.62	Random	200	5	5	2.64
Noise+100	Random	200	2	5	2.99	Random	200	2	5	3.04

Table III

Input Code Letter	Description of Input	
	Increasing Ramp, Ft/Sec	Decreasing Ramp, Ft/Sec
A	Noise+20	20
B	Noise+20	80
C	Noise+80	20
D	Noise+80	80

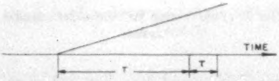


Fig. 8. Ramp input adaptive system

The evaluation of the adaptive system may be carried a step further by comparing its performance with the performance of the optimum filter that will reproduce a ramp input when contaminated with white noise. The weighting function for such a filter may be determined by the method of Zadeh and Ragazzini,[3] and is:

$$W(t)=\frac{4}{T}-\frac{6}{T^2}\,t \qquad 0\leq t\leq T$$

$$=0 \text{ elsewhere.}$$

The mean square noise output σ_o^2 of such a filter, when subjected to the same noise input as the adaptive system is

$$\sigma_o^2=\frac{1.272}{T}$$

Inspection of the above expression shows that values of T greater than about 1/2 second are necessary in order that the optimum filter have a root-mean-square noise output comparable to the adaptive system. For such values of T, the transient recovery period of the optimum filter is considerably longer than that of the adaptive system. In addition, since the weighting function of the optimum filter is discontinuous at time T, it is not realizable in terms of lumped elements. Thus an approximate weighting function must be determined which is then realized in terms of lumped elements.

VI. Analog Computer Simulation of Adaptive Servo System

In order to evaluate the performance of the adaptive servo, the example of the preceding section was simulated, along with two linear systems, on an Electronics Associates analog computer. The linear systems were chosen as the two extremes: the one to represent a good smoothing system with correspondingly large time constants, and the other a good tracking system with small time constants. All three were subjected to the same inputs, and records of the output errors were obtained.

More particularly, the two linear systems were

1. Good tracker

$\omega_n=2\pi\times 8$ radians per second, $\zeta=0.5$

2. Good smoother

$\omega_n=2\pi\times 1$ radians per second, $\zeta=0.5$

Because of equipment limitations it was possible only to simulate the adaptive system, together with one linear system, in each sequence of tests.

The simulated systems were subjected to the following inputs: (1) noise plus ramps, as shown in Fig. 8; and (2) noise plus an increasing ramp followed by a decreasing ramp, as shown in Fig. 9.

The noise signal was derived from a noise generator with a spectral density that was essentially constant (adjusted to 0.636 square foot per second) up to a frequency of 60 cycles per second.

In accordance with the objective of the simulation (to provide a quantitative comparison between adaptive and linear systems), the error signal, that is, the

Table IV

Input	Adaptive System Errors, Feet			Linear Tracker Errors, Feet		
	ϵ_L*	ϵ_C*	ϵ_R*	ϵ_L	ϵ_C	ϵ_R
A	2.66	2.28	2.34	2.75	2.67	3.07
B	2.41	2.78	3.37	2.41	2.70	2.92
C	3.63	2.98	2.47	2.84	2.63	2.19
D	3.34	3.06	3.52	3.22	2.84	2.98

* ϵ_L, ϵ_C, and ϵ_R are the rms errors in tracking combination of ramp inputs, as shown in Fig. 9, over intervals τ_L, τ_C, and τ_R, respectively.

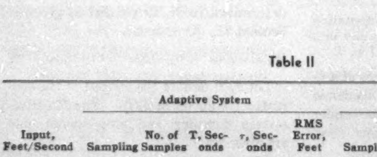


Fig. 9 (left). Combination ramp input to adaptive system

Table II

Input, Feet/Second	Adaptive System					Linear Smoother				
	Sampling	No. of Samples	T, Seconds	τ, Seconds	RMS Error, Feet	Sampling	No. of Samples	T, Seconds	τ, Seconds	RMS Error, Feet
Noise	Uniform	200		9.25	1.46	Uniform	200		9.25	1.03
Noise+20	Random	200	5.63	5	2.57	Uniform	81	5.63	5	3.87
Noise+50	Random	200	1.5	5	3.18	Uniform	81	1.5	5	8.91
Noise+100	Random	100	3	2.5	3.63	Uniform	49	2.5	3	17.00

Table V

Input	Adaptive System Errors, Feet			Linear Smoother Errors, Feet		
	ϵ_L	ϵ_C	ϵ_R	ϵ_L	ϵ_C	ϵ_R
A	2.08	1.88	2.92	3.71	3.49	3.81
B	2.35	2.42	3.15	3.57	8.23	11.30
C	3.22	1.72	1.85	8.55	3.68	3.58
D	3.62	3.36	3.91	14.3	10.03	13.9

difference between the servo output and the noiseless ramp input, or combination of ramp inputs, was recorded in each case. Since the error signal is a random function, it was sampled either in uniform intervals of τ seconds or, in some cases, at random intervals; this fact is noted in the tabulated results. Time T, which was allowed for transients to die out, is also noted.

Table I, which contains the comparison between the adaptive servo and the good linear tracker, shows not only that the former smooths noise in the absence of a signal better than the latter by a factor of two, but also shows that it performs equally well in tracking ramps through noise. Table II exhibits corresponding results for the adaptive servo compared with the good linear smoother. The smoothing properties of the two are here shown to be about the same in the case of noise only; in tracking a ramp through noise, however, the adaptive servo is quite superior.

Similar results for combinations of ramp inputs are contained in Tables III, IV, and V.

In order to evaluate qualitatively the transient behavior of the adaptive system, responses to step inputs were simulated. At the outset it was feared that, since the design procedure was aimed chiefly at good steady-state behavior, the transient response might develop to be poor. The opposite was found to be true. The recovery time of the servo appeared to be substantially shorter than that of even the good linear tracker and exhibited a damping ratio equivalent to about $\zeta = 0.5$.

Appendix

Weighting Function of Optimum Velocity Estimator

Following the procedure derived by Zadeh and Ragazzini,[3] let $W(t)$ be the weighting function of the estimator

$$W(t) = 0 \text{ for } t<0 \text{ and } t>T,$$

and let

$$\mu_i = \int_0^\infty dt\, t^i W(t) = \int_0^T dt\, t^i W(t)$$

be the ith moment of $W(t)$. Let the estimator be subjected to a signal consisting of a polynomial $P(t)$ of, at most, the second degree and a random noise with autocorrelation.

$$R_N(\tau) = (1/2)\sigma_N^2 \delta(\tau)$$

and spectral density

$$G_N(f) = \sigma_N^2$$

For a velocity estimator

$$\mu_0 = 0 \qquad \mu_1 = -1 \qquad \mu_2 = 0$$

The weighting function is given by

$$\begin{aligned} W(t) &= A_0 + A_1 t + A_2 t^2 \qquad 0 \le t \le T \\ &= 0 \text{ elsewhere} \end{aligned}$$

where A_0, A_1, and A_2 are constants determined from the moments of $W(t)$. Thus,

$$A_0 T + \frac{A_1}{2} T^2 + \frac{A_2}{3} T^3 = \mu_0$$

$$\frac{A_0}{2} T^2 + \frac{A_1}{3} T^3 + \frac{A_2}{4} T^4 = \mu_1$$

$$\frac{A_0}{3} T^3 + \frac{A_1}{4} T^4 + \frac{A_2}{5} T^5 = \mu_2$$

Solving the above equations, we obtain

$$A_0 = \frac{36}{T^2}, \quad A_1 = \frac{192}{T^3}, \quad A_2 = \frac{180}{T^4}$$

Hence,

$$\begin{aligned} W(t) &= \frac{36}{T^2} - \frac{192}{T^3} t + \frac{180}{T^4} t^2 \qquad 0 \le t \le T \\ &= 0 \text{ elsewhere} \end{aligned}$$

Choice of the Estimator Time Constant

If the estimator is subjected to an input consisting of a noise $N(t)$ and a polynomial of the form

$$P(t) = A_0 + A_1 t$$

The steady-state output will consist of a noise with root-mean-square value σ_E superposed on the derivative of $P(t)$, i.e., A_1.

Let α_1 be the signal-to-noise ratio at the estimator output. Then,

$$\alpha_1^2 = \frac{A_1^2}{\sigma_E^2} = \frac{T^3}{96} \frac{A_1^2}{\sigma_N^2}$$

or

$$T = \left[\frac{96 \alpha^2 \sigma_N^2}{A_1^2} \right]^{1/3}$$

Thus if α_1 is the minimum acceptable signal-to-noise ratio for the estimator, then A_1 is the smallest value of velocity input that can be detected. The estimator time constant T is then given by the above expression.

References

1. Theory of Servomechanisms, H. M. James, N. B. Nichols, R. S. Phillips. *Radiation Laboratory Series*, Massachusetts Institute of Technology, Cambridge, Mass., vol. 25.

2. An Adaptive Servo System, A. H. Benner, R. F. Drenick. *Convention Record*, Institute of Radio Engineers, New York, N. Y., pt. 4, 1955, p. 8.

3. An Extension of Wiener's Theory of Prediction, L. A. Zadeh, J. R. Ragazzini. *Journal of Applied Physics*, New York, N. Y., vol. 21, July 1950.

◆

Discussion

Ian McCausland (University of Toronto, Toronto, Ont., Canada): The authors define an adaptive control system as one which monitors its own performance and adjusts some of its own parameters in the direction of better performance. However, any feedback control system may be said to monitor its own output or outputs and to adjust something in the system in such a way as to decrease the discrepancy between what the performance is and what it ought to be. To call one system adaptive because it varies parameters, and to exclude another system from this category because it only varies signals, seems to be an undue restriction of the meaning of the word "adaptive." Is the variation of parameters the essential feature of an adaptive system? If so, an ON-OFF servomechanism is adaptive since it varies the loop gain from zero to a definite positive or negative value in accordance with fixed rules. If we include an ON-OFF servomechanism in the adaptive category can we exclude the proportional-control system, which provides a continuously-variable adjustment of the gain between the power-supply and the load?

The authors state that the term "adaptive" has been borrowed from biology. However, in the science of biology this term describes a number of different types of process. For example, Ashby[1] proposes the definition that "a form of behavior is adaptive if it maintains the essential variables within physiological limits"; and many of the examples given by him to illustrate adaptive behaviour are not essentially different from the conventional feedback control system. The meaning of the term "adaptation" in the theatrical field seems to be closer to what the authors have in mind, that is, the alteration of a piece of dramatic writing in order to suit the taste of a different audience or to accommodate it to different circumstances.

Since interest in the type of servomechanism described by the authors is likely to increase in the future, it appears to be necessary at this stage to choose technical terms which are as precise as possible. Since the term "adaptive" is used so widely in biology it seems desirable to qualify it in some way, so that it will not be taken to describe the entire class of feedback control systems.

REFERENCE

1. Design for a Brain (book), W. Ross Ashby. John Wiley & Sons, Inc., New York, N. Y., 1954, chap. 5.

R. F. Drenick and R. A. Shahbender: Mr. McCausland raises a point of terminology. He agrees that control systems of the type discussed by us belong to a class of such systems and that they differ from the conventional type. However, he feels that the term "adaptive" is overly comprehensive and therefore a poor choice.

Points of terminology are always difficult to settle but the first thing to do in any case is to try to define the class of systems to which a term is to be attached. As far as we know, this has not been done. We shall therefore make an attempt at it. It will of necessity be a rather broad and unspecific definition, partly because the state of our knowledge about these systems, and partly because the nature of this discussion, preclude going into much detail. Nevertheless, it will become quite clear that the class as we hope to define it will include the servo mechanisms which we have discussed, along with many others; that it does not include the ON-OFF system or the proportional-control system mentioned in his comments. Finally, we hope to make plausible our choice of the term "adaptive" for the class of systems.

Broadly speaking, we expect the servo mechanisms of this class to be designed for operation in a slowly changing environment, as opposed to the conventional types which are intended for a fixed environment. To be more specific, we must define what we mean by "environment." We use the term here collectively for all the data, given or assumed, which must be taken into account by the system designer, that is:

1. The nature of the signal for which the system is to be designed.
2. The nature of the noise (if any) against which it should discriminate.
3. The set of parameters (if any) of the system which are given at the outset and which cannot be negotiated by the designer.

It is correct to say, we believe, that a conventional servomechanism is intended for operation in a fixed environment, or in any case, is designed with such an environment in mind. That is to say, the nature of the signal and of the noise (their spectral densities, say) and of the nonnegotiable parameters (some of the gains and time constants) are fixed or, if necessary, are artificially frozen at some reasonable values. The control system is then so designed that it will operate in this environment in some optimal, or nearly optimal, fashion.

We propose that, in order for a control system to belong to the class which we have called adaptive, two conditions should be fulfilled, namely:

1. The environment should be allowed to change slowly (when compared with the frequencies of signal and noise), the system should be able to detect changes and vary its (negotiable) parameters accordingly.
2. The system should have a single input, comprising signal and possibly noise, and it should be designed to derive all indications of environmental change either from that input, or from its own response to it, or both.

We omit discussion of these specifications or of their motivation. It is probably worth pointing out that the class is rather broad, that it includes the type of system we dealt with in our paper, and many others, but that it does not include ON-OFF servos nor conventional linear servos (except possibly as a pointless border line case).

We can now comment briefly on the matter of terminology. To our knowledge, the following terms have been attached to control systems essentially of the class defined here: "input-controlled," "sympathetic," "learning," "self-adjusting," "self-designing," and "adaptive." The last, which was our choice, is used in biological fields in the sense of "modification of an animal or plant (or its parts or organs) fitting it more perfectly for existence under the condition of its environment," according to Webster's Collegiate Dictionary. We felt, and still feel, that this describes very well the control systems under discussion.

Quasi-Optimization of Relay Servos by Use of Discontinuous Damping

W. L. HARRIS, JR.
STUDENT MEMBER AIEE

C. McDONALD
STUDENT MEMBER AIEE

G. J. THALER
MEMBER AIEE

A NUMBER of articles[1-4] have been written which deal with the optimum response of a relay servo to a step input, with some mention of a ramp input. If optimum response is defined as that response which regains position correspondence in minimum time and without overshoot, then it has been shown that an ideal relay servo with second order motor-load combination may be optimized. The optimized relay servo has one period of acceleration with maximum torque, an instantaneous torque reversal (ideal relay) at an appropriate point, and one period of deceleration with maximum torque. The system attains zero error and zero error-rate simultaneously, the torque is removed, and the system is stationary at zero error. For the second order motor-load, with a maximum torque specification, this represents the fastest possible response. For higher order motor-load units, more switching operations[5] are required.

In order to realize such an optimized servo, an anticipating device is required to reverse the torque at the proper instant. Phase plane analysis shows that the locus of points on the phase plane at which this reversal is required is precisely that deceleration trajectory which passes through the origin of the phase plane. This is shown in Fig. 1. Physical realization of this switching locus requires a nonlinear anticipator. Satisfactory operation of the servo depends on the accuracy with which the nonlinear device produces the desired switching locus, and also depends on the properties of the relaying device, which has been postulated as ideal.

A quasi-optimum relay servo is hereby defined as one which has all of the characteristics of an optimum relay servo except that of minimum response time. In practice, an ideal relay does not exist. Assuming that the only defect in the relay is the existence of a dead zone, a quasi-optimum relay servo may be obtained by the application of a damping or braking torque when the relay opens. If the damping is properly controlled, the system comes to rest at or near zero error, and it will be shown that this can be accomplished for a wide range of step inputs. The resulting system is deadbeat, has one accelerating period under full torque, and one decelerating period with variable torque as a result of the natural damping plus the discontinuously applied damping.

Nomenclature

C=coulomb friction torque in pound-feet
D=added viscous damping torque
$E=\theta_R-\theta_c$=position error in radians
E_{ss}=steady-state position error in radians
f=coefficient of viscous friction in pound-feet per radian per second
I_a=motor armature current in amperes
J=moment of inertia in slug-feet squared
K_t=motor torque constant in pound-feet per ampere
K_v=motor back emf constant in volts per radian per second
L_a=armature inductance in henries
R_a=armature resistance in ohms
R_B=dynamic braking resistance in ohms
T=motor shaft torque in pound-feet
V=voltage applied through relay in volts
θ_c=output shaft position in radians
θ_R=reference shaft position in radians

Paper 57-605, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Great Lakes District Meeting, Des Moines, Iowa, April 15–17, 1957. Manuscript submitted January 8, 1957; made available for printing May 17, 1957.

W. L. HARRIS, JR., C. MCDONALD, and G. J. THALER are with the United States Naval Postgraduate School, Monterey, Calif.

The opinions contained herein are the private ones of the authors and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large.

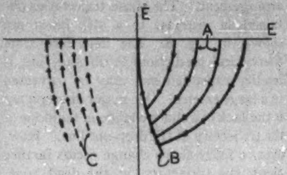


Fig. 1. Phase trajectories for an optimum relay servo

A—Acceleration trajectories
B—Deceleration trajectories
C—Deceleration trajectory and locus of relay reversals

Phase Plane Analysis of Discontinuous Damping

For the purpose of introducing the basic theory, it is assumed that the relay servo consists of an ideal error detector, pure torque motor, and pure inertia load. The basic differential equation during acceleration periods is

$$J\ddot{E}=+T \tag{1}$$

If the relay has dead zone, when the relay opens the differential equation is

$$J\ddot{E}=0 \tag{2}$$

The phase trajectory corresponding to these conditions is sketched in Fig. 2(A). Note that the system traverses the dead zone at constant velocity. If the motor-load unit contains some viscous damping, then equations 1 and 2 are altered and the effect on the phase trajectory is shown by dotted curve of Fig. 2(A). Note that the damping causes some deceleration in the dead zone. If some damping is added to the output at the instant the relay opens, then the output is decelerated as it crosses the dead zone. Furthermore, if the added damping is viscous damping, then equation 2 becomes

$$J\ddot{E}=-D=-f\dot{E} \tag{3}$$

Such discontinuous introduction of damping does not affect equation 1, and therefore permits maximum acceleration when the motor is driving. The resultant phase trajectory is shown on Fig. 2(B). Note that the deceleration trajectory in the dead zone is a straight line. This is a convenient and fortuitous circumstance. If an appreciable amount of coulomb friction is present, this deceleration trajectory is curved.

Addition of a linear differentiating device, either in the error channel or in feedback, rotates the dead zone boundaries. Such a condition is illustrated in Fig. 2(C). Note that this provides anticipatory switching according to a linear law. From the phase trajectory of Fig. 2(C), the effect is obviously stabilizing. If the magnitude of the coefficient of the discontinuous viscous damping is considered in determining the rotation to be produced by the differentiating device, the phase trajectory in the dead zone can be made precisely parallel to the dead zone boundary, and the trajectory falls exactly on this boundary, as shown in Fig. 2(D). The system is then deadbeat for all magnitudes of step input. The response time depends on the magnitude of the viscous damping, as well as the magnitude of the step input, and increasing the damping coefficient permits reduction in response time.

The deadbeat response shown in Fig. 2(D) has a steady state error E_{ss} which is equal to one half of the total dead zone width. However, since the trajectory does not enter the dead zone but merely follows its boundary, the width of the dead zone may be decreased until the dead zone width is infinitesimal and the steady state error is negligible, i.e., $E_{ss}\cong 0$. This has been illustrated on Fig. 2(E). This, then, represents a quasi-optimized condition.

Note that if the viscous damping is made sufficiently large, the dead zone approaches a vertical position and the response time may be even shorter for the quasi-optimized system then for the so-called optimum relay servo.

Discontinuous Damping by Dynamic Braking of a D-C Shunt Motor

The theory presented in the preceding section is readily applied to a relay servo using a d-c shunt motor. Discontinuous damping is obtained by connecting a dynamic braking resistor across the armature terminals of the motor, as shown in Fig. 3. Note that the only extraneous effect during acceleration intervals is to draw additional current through the relay contacts, but this does not affect the dynamic performance of the motor.

Assuming that the armature inductance and the coulomb friction are negligible and that the relay is ideal, it may be shown that the phase plane equations are

$$\frac{d\dot{E}}{dE}=-\frac{f}{J}-\frac{K_TK_e}{JR}-\frac{K_TV}{JR\dot{E}} \tag{4}$$

for the acceleration period, and

$$\frac{d\dot{E}}{dE}=-\frac{f}{J}-\frac{K_TK_e}{J(R_aR_B)} \tag{5}$$

for the deceleration in the dead zone. (The symbols are defined in the nomenclature section, and the equations are derived in Appendix I.) Equations 4 and 5 define the isoclines on the phase plane. Equation 4 is typical of torque saturated operation and needs no comment, but equation 5 shows that the trajectory in the dead zone is a straight line, and that the slope of this line may be adjusted by varying R_B. Since these equations show that the d-c shunt motor with dynamic

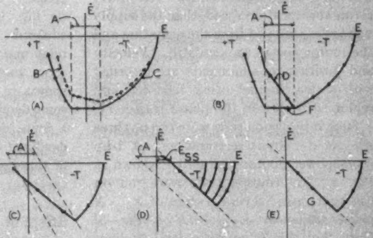


Fig. 2. Phase trajectories describing the effect of discontinuous damping

A—Trajectories of basic relay servo
B—Effect of discontinuous damping
C—Rotation of dead zone by differentiator
D—Matching the slopes of trajectory and dead zone boundary
E—Reducing width of dead zone to optimize response

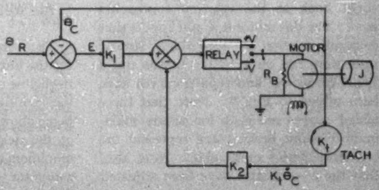


Fig. 3. Block diagram of a shunt motor relay servo with dynamic braking for discontinuous damping

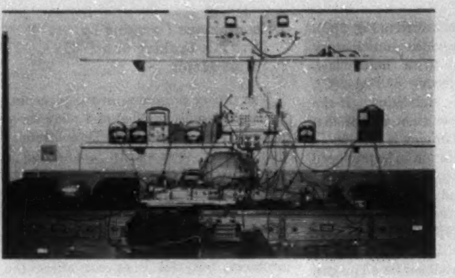


Fig. 4. The 1-hp motor load unit with associated control and measurement devices

braking has precisely the phase plane characteristics required by the theory, actual performance will vary from the theoretical only if the assumptions are not valid. Experimental verification of the theory is, therefore, necessary.

Characteristics of D-C Servomotors, Coulomb Friction Effects

It is obvious that relays in general have some time lag in operation, motor armatures possess inductance,[7] and motor-load combinations have some coulomb friction. Thus, it is known that the three basic assumptions are not completely valid. However, if these effects are negligible, or if they may be minimized by proper design, the resulting system performance can be satisfactory The coulomb friction effect is readily checked prior to assembling the servo. The motor, dynamic braking resistor, gear train, tachometer, and error detector are assembled as they are to be used in the servo, thus, including all friction effects. The motor is run at saturation speed, then the supply voltage is removed from the armature and the motor coasts to standstill. Velocity and position measurements are recorded with a two channel Brush recorder, and replotted to obtain the phase trajectory.

Such retardation tests were run on three sizes of d-c shunt servomotors. A 1-hp (horsepower) motor-load unit is shown in Fig. 4. Retardation tests were run on this unit, using several values of R_B. A phase trajectory is plotted on Fig. 5. Note that this is nearly a straight line; the only perceptible coulomb effect is the slight hook as the velocity approaches zero. For this motor it would appear that coulomb friction is negligible. Similar tests were run on a 1/8-hp motor and a 1/125-hp motor, and typical curves have been added to Fig. 5. Note that these added curves are each for nearly maximum dynamic braking and represent the best trajectories obtainable. Note also, that the co-ordinates have been adjusted by proportion to facilitate comparison of the trajectories' curvatures. From Fig. 5 it is seen that the largest motor shows the least coulomb friction effect, its trajectory being an almost perfect straight line. Also, the smallest motor shows the greatest coulomb friction effect, its trajectory being considerably curved. It can be shown mathematically that this variation in curvature is largely due to the C/J ratio; see Appendix II.

For all of the cases shown in Fig. 5, the coulomb friction effect would prevent the phase trajectory from following the dead-zone boundary exactly. However, practical relays have a pull-in point which requires more voltage than is needed to hold the relay closed, so that the practical relay drops out at a lower voltage then the pull-in voltage. The phase portrait therefore requires' additional boundary lines for the dead zone, as shown in Fig. 6(A), and for slight curvature of the deceleration trajectory the system is completely deadbeat, but with slightly increased static error, as also shown on Fig. 6(A). For larger amounts of coulomb friction, the forward driving relay may close several times producing a jerky motion. The phase trajectory is shown in Fig. 6(B). The net result is some decrease in the static accuracy, since a finite dead zone is required to obtain deadbeat operation. The actual width of this dead zone depends on the specific system.

Effect of Relay Time Lag and Armature Inductance as Shown by Step Function Tests on the Relay Servo

The 1-hp motor-load unit was selected for the initial investigation because of the negligible coulomb friction effect. It was incorporated in a relay servo, the circuit being shown in Fig. 7. Note that control of the dead zone was easy because the amplifiers used were taken from an analog computer and provided a very flexible arrangement. The phase trajectories obtained in response to a step input are shown in Fig. 8. The departure from theoretical predictions is obvious, but is readily explained and may be corrected to a large extent. One major discrepancy is the lack of a sharp change in the slope of the trajectory at the drop-out line. However, a fairly sharp change occurs farther along the trajectory, in the dead zone. This suggests a time lag in relay operation, and auxiliary tests verified this. The importance of minimizing relay lag thus becomes obvious. A second discrepancy in the trajectories of Fig. 8 lies in the fact that the trajectory does not become a straight line immediately after the relay opens, but continues curving before it straightens out. This is caused by armature inductance. Fig. 9 shows a curve of armature current versus time. The relay closes at *A*, and the current rises rapidly (but not instantaneously) to a maximum, then decays along *B* to a nonzero steady state at *C*. When the relay drops out, the current must first decrease to zero, then reverse, going through a maximum along the curve *CD*. Note that a finite time is required for the current to decrease to zero, reverse, and rise to a maximum in the reverse direction. During this time interval, the dynamic braking effect is not as large as predicted, and the phase trajectory curves out into the dead zone. For smaller disturbances, the relay drops out earlier, i.e., nearer *B* on Fig. 9, and the effect of the armature inductance has a greater duration because it takes longer to

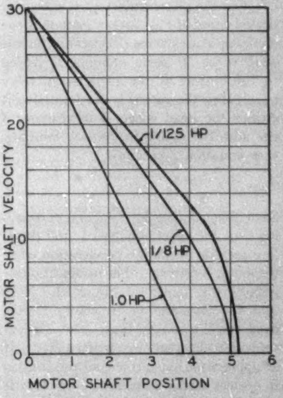


Fig. 5. Effect of coulomb friction on the phase trajectories of dynamically braked d-c shunt motors

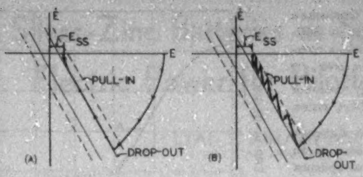


Fig. 6. Effect of coulomb friction on a relay servo when the relay has unequal pull-in and drop-out voltages

A—Effect of a small amount of coulomb friction
B—Effect of a large amount of coulomb friction

reduce the initial armature current to zero.

It is immediately obvious that the relay time lag can be decreased by using a better relay, and the delay due to armature inductance can be reduced by selecting a motor such that the electrical time constant of the armature, L_a/R, is small. Another factor to be noted is that both effects may be compressed on the phase plane by increasing the motor-load gear ratio. The wide dead zone of Fig. 8 is largely due to the fact that the gear ratio was chosen to emphasize the defects, and was only 13 to 1, so that an input step of 25 degrees was corrected in less than one revolution of the motor shaft.

Step and Ramp Function Tests on the Relay Servo Using a 1/125-Hp Motor

In order to permit use of a smaller and faster relay, the 1/125-hp motor replaced the larger motor. This motor also had a much more favorable electrical time constant for its armature, and finally the gear ratio was raised to 500 to 1. The phase trajectories for the response of this system to a step input are shown in Fig. 10. It is readily seen that these trajectories more nearly approximate those predicted by theory, and the width of the dead zone has been greatly reduced. Still further improvements can be made by means of motor selection, relay selection, and gear ratio adjustment.

This system was also tested with a ramp input. Due to the use of tachometer feedback, a velocity lag error is introduced. Substitution of an error differentiator eliminated the velocity lag error. The use of a ramp input always causes a limit cycle, which is characterized by opening and closing of the forward driving relay only; the reverse relay never operates. It is thus seen that the output during ramp operation is essentially confined to the region between the pull-in and the drop-out points. The frequency of the limit cycle increases with increased input velocities.

Comparison of a Quasi-Optimized Servo with an Optimum Servo Using the Same Motor-Load Unit

Fig. 11(A) shows an experimentally obtained phase trajectory for the 1/125-hp quasi-optimized servo, and also an optimum phase trajectory for the same motor-load unit obtained from experimental acceleration and decleration trajectories. There is an appreciable difference between these trajectories in the deceleration region, so Fig. 11(B) shows the transient response curves obtained from these trajectories. It is readily seen that the responses are identical during acceleration, nearly identical for part of the deceleration period, but the quasi-optimized system approches correspondence more slowly when the error has been reduced to a small value.

Conclusions

A theory has been developed for using discontinuous damping to obtain a relay servo which has a deadbeat response to a step input. This theory has been applied to the case of a relay servo using a d-c shunt motor, in which the discontinuous damping is obtained by connecting a dynamic braking resistor across the motor armature. It has been shown that static accuracy is adversely affected by relay time lag, by the electrical time constant of the armature, and by coulomb friction, but it has also been shown that these effects can be controlled and an acceptable system obtained. It is obvious that the quasi-optimized system is simpler to design, construct, and adjust than a comparable system using optimum switching, since the latter requires a special nonlinear anticipator for switching. It has also been shown that the transient response of the quasi-optimized system is not too unfavorable as compared with the optimum response. Finally, it might be noted that the dynamic braking resistor tends to lock the output and prevent wander in the dead zone when the system is at rest.

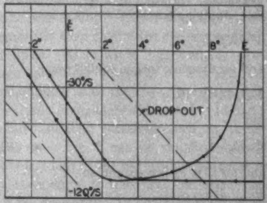


Fig. 8. Phase trajectories of a quasi-optimized relay servo using a 1-hp shunt motor (step response)

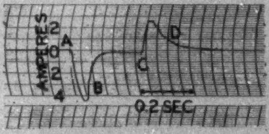


Fig. 9. Armature current variation in a dynamically braked relay servo

AB—Initial inrush
BC—Decrease to steady state as motor accelerates
CD—Reverse current due to dynamic braking

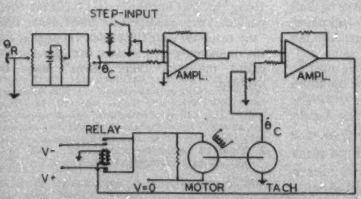


Fig. 7 (left). Circuit diagram of quasi-optimized relay servo using 1/125-hp d-c shunt motor

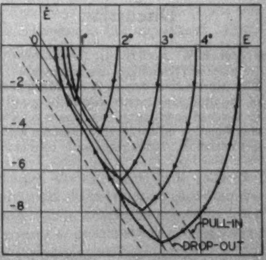


Fig. 10 (right). Phase trajectories showing the step response of a quasi-optimized relay servo using a 1/125-hp shunt motor

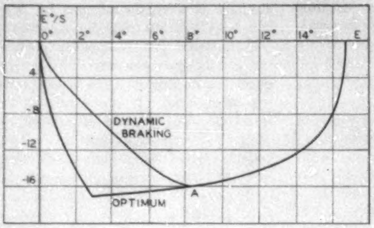
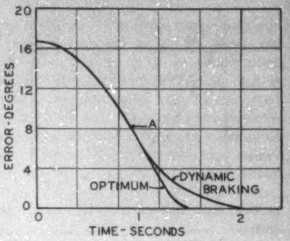


Fig. 11(A) (left). Comparison of phase trajectories for optimum and quasi-optimum operation of a relay servo

A—Point of relay drop-out for the quasi-optimum system

Fig. 11(B) (right). Comparison of time response of an optimum and a quasi-optimum relay servo

A—Point of relay drop-out for the quasi-optimum system



Appendix I

Derivation of Equations 4 and 5

Assumptions: C/J is negligible; $4R_a$ is negligible; the input is a step of position, θ_R. Applying Kirchhoff's voltage law,

$$V=K_v\dot{\theta}_c+I_aR_a \tag{6}$$

and for dynamic torque equilibrium,

$$K_tI_a=J\ddot{\theta}_c+f\dot{\theta}_c \tag{7}$$

but

$$E=\theta_R-\dot{\theta}_c \tag{8}$$

and for $t>0$, θ_R is constant, so

$$\left.\begin{aligned}\dot{E}&=-\dot{\theta}_c \quad (t,0)\\ \ddot{E}&=-\ddot{\theta}_c \quad (t,0)\end{aligned}\right\} \tag{9}$$

From which

$$V=-K_v\dot{E}+I_aR_a \tag{10}$$

$$K_tI_a=-J\ddot{E}-f\dot{E} \tag{11}$$

which may be combined to give the following:

$$\ddot{E}=-\frac{K_tV}{JR_a}-\frac{K_tK_v\dot{E}}{JR_a}-\frac{f}{J}\dot{E} \tag{12}$$

Dividing both sides by $dE/dt=\dot{E}$,

$$\frac{\frac{d^2F}{dt^2}}{\frac{dE}{dt}}=\frac{\frac{d\dot{E}}{dt}}{\frac{dE}{dt}}=\frac{d\dot{E}}{dE}=\frac{-K_tV}{JR_a\dot{E}}-\frac{K_tK_v}{JR_a}-\frac{f}{J} \tag{4}$$

which is the equation for the isoclines during the accelerating period.

For the deceleration in the dead zone due to dynamic braking, the following equations apply:

$$-K_v\dot{E}+I_aR_a+I_aR_B=0 \tag{13}$$

$$K_tI_a=-J\ddot{E}-f\dot{E} \tag{14}$$

These combine to give:

$$\ddot{E}=\frac{-K_tK_v\dot{E}}{J(R_a+R_B)}-\frac{f\dot{E}}{J} \tag{15}$$

Again dividing both sides by $dE/dt=\dot{E}$

$$\frac{\frac{d\dot{E}}{dt}}{\frac{dE}{dt}}=\frac{d\dot{E}}{dE}=\frac{-K_tK_v}{J(R_a+R_B)}-\frac{f}{J} \tag{5}$$

which is the equation for the isoclines in the dead zone.

Appendix II

Effect of Coulomb Friction on the Deceleration Trajectory

In this case, equation 13 is valid, but equation 14 must include a coulomb friction term, and becomes:

$$K_tI_a+J\ddot{E}+f\dot{E}-C=0 \tag{16}$$

Combining equations 13 and 16,

$$\ddot{E}=-\frac{K_tK_v}{J(R_a+R_B)}\dot{E}-\frac{f}{J}\dot{E}+\frac{C}{J} \tag{17}$$

Dividing both sides by $dE/dt=\dot{E}$,

$$\frac{\frac{d\dot{E}}{dt}}{\frac{dE}{dt}}=\frac{d\dot{E}}{dE}=-\frac{K_tK_v}{J(R_a+R_B)}-\frac{f}{J}+\frac{C}{J\dot{E}} \tag{18}$$

This equation for the isoclines shows that the phase trajectory curves as $\dot{E}$ varies, providing $C/J\neq 0$.

References

1. A Phase-Plane Approach to the Compensation of Saturating Servomechanisms, Arthur M. Hopkin. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

2. Nonlinear Techniques for Improving Servo Performance, D. McDonald. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 400–21.

3. A Study of Some Nearly Optimum Servomechanisms, T. M. Stout. *Report no. 10*, Electrical Engineering Department, University of Washington, Seattle, Wash., 1952.

4. Optimization of Nonlinear Control Systems by Means of Nonlinear Feedbacks, R. S. Neiswander, R. H. MacNeal. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 262–72.

5. An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms, Irving Bogner, Louis F. Kazda. *Ibid.*, vol. 73, pt. II, July 1954, pp. 118–27.

6. Optimum Switching Criteria for Higher Order Contactor Servo With Interrupted Circuits, S. S. L. Chang. *Ibid.*, vol. 74, pt. II, Nov. 1955, pp. 273–76.

7. Transfer Function and Parameter Evaluation for D-C Servomotors, G. J. Thaler, W. A Stein. *Ibid.* (Jan. 1956 section), pp. 410–17.

8. Discontinuous Damping of Relay Servomechanisms, W. L. Harris, Jr. *M.S. Thesis*, U. S. Naval Postgraduate School, Monterey, Calif., 1956.

◆

Discussion

Frederick W. Nesline, Jr. (Massachusetts Institute of Technology, Cambridge, Mass.): The paper's contribution is the adjustment of the dead-zone trajectories to be parallel to the dead-zone boundaries so as to achieve step function or disturbance response without overshoot. This response is then compared to that of an optimum relay servomechanism. Such a comparison is not a good one because the constraints necessary to derive the optimum system used for comparison are violated by the quasi-optimum system. It is more logical to compare the quasi-optimum system with the optimum system that has the same constraints, i.e., discontinuous damping must be allowed in both cases. This true optimum does have a vertical phase trajectory along the $\dot{E}$ axis and exhibits the property of having the fastest possible time response to certain inputs for this set of constraints. Using the true optimum as a basis for comparison, the paradoxical statement that "the response time may even be shorter for the quasi-optimized system than for the so-called optimum relay servo" may be avoided.

Silver-Zinc Batteries as Source-Primary Electric Power for Pilotless Aircraft

LYAL T. ABRAHAMSON
ASSOCIATE MEMBER AIEE

THE design of a pilotless aircraft electric system must be approached in a far different manner than for a piloted airplane. The pilotless aircraft has nobody to substitute for inoperative motors, relays, or servo systems. Each component in the aircraft system must operate properly or the result very likely is an aborted mission. Also, at the speeds and accelerations encountered in pilotless aircraft, a tremendous amount of energy is expended to get each pound of equipment to its destination, thus an ever-increasing premium is placed on component weight. To achieve the conflicting requirements of high reliability and light weight with equipments either available or within reach of reasonable development time and efforts, the system depicted by Fig. 1 was designed and developed.

Note that a single-string system is used to keep weight to a minimum. This scheme depends on exceptionally high reliability in each component right down to the last nut, screw, and solder joint. It also follows that fuses are not required where single power sources supply systems which are essential to successful mission completion.

General System Description

A battery, power contactor, d-c motor, and variable displacement pump furnish hydraulic power for movement of the control surfaces. The battery is made up of a single string of series-connected secondary silver-zinc cells. Circuit redundancy is provided by multiterminal cells and parallel wires in the connecting cables. The d-c compound wound motor is directly connected to external connectors through which power is provided for manufacturing tests and prelaunch checkouts. Just prior to launching, the battery is connected to the motor by actuating the power contactor.

Power for the guidance and telemeter systems is furnished by three electrically separate battery supplies which are the +6.75-volt filament battery, the +28-volt inverter battery, and the multivoltage high-voltage battery. In addition to the 400-cycle-per-second 115/200-volt inverter, the heat-exchanger fan motor, small relays, and motors are supplied by the +28-volt battery. All three batteries are made up of single strings of series-connected secondary silver-zinc cells. Circuit redundancy is limited to the use of parallel pins in miniature connectors. As in the hydraulic system, aircraft checkouts are performed on externally supplied power. Prior to launching, the batteries are connected to their respective loads by the power changeover switch and power contactors.

Except in cases where power source isolation is required, all regulated power is provided by centralized voltage regulators. The (+250 ± 1%)-volt and (+150 ± 1%)-volt regulators are supplied by the high-voltage battery, and the (−250 ± 1%)-volt regulator derives its primary power from the inverter.

Silver-Zinc Batteries

CONSTRUCTION AND CHEMISTRY

The silver-zinc electrochemical system, although known for many years to be a highly concentrated source of electric power, was impractical until recently because of the very limited storage life. Discovery of an ion-selective separator material and subsequent improvements to it have made possible cells of the type utilized in the power system shown in Fig. 1.

In the discharged state the positive plates are made up of porous silver which is mechanically strong enough to maintain its original shape even when removed from the cell. The cathodes consist of zinc hydroxide powder contained in paper sacks. The electrolyte is a 40% solution of potassium hydroxide and water. The separator material surrounds the silver plates and prevents migration of silver ions to the cathode. The Bakelite *C-11* cases are designed to allow leak-free operation in any position, but in storage for extended periods should be upright.

A simplified over-all cell reaction is given by the following equation:[1]

$$AgO+Zn+H_2O \underset{\text{Discharged}}{\overset{\text{Charged}}{\rightleftarrows}} Zn(OH)_2+Ag$$

The reaction actually involves an intermediate state of the positive plate. As the charging begins the silver is converted to silver oxide (Ag_2O), as evidenced by a cell voltage of approximately 1.65 volts. When some of the silver oxide is converted to silver peroxide (AgO), the cell voltage increases to approximately 1.90 volts. The voltage rises above 2.05 volts when the cell is fully charged and oxygen evolution begins at the positive plate.

GENERAL CHARACTERISTICS

In addition to providing a high concentration of potential energy, the silver-zinc cells exhibit a flatter discharge characteristic than most other electrochemical couples. As shown in Figs. 2 and 3, voltage tolerances of ±6% or less can be maintained until the capacity of the cell is nearly completely exhausted. In fact, moderate load increases such as those shown in Fig. 3 can be supplied during the last half of the battery discharge at a voltage equal to that near the beginning of the discharge. In contrast, the nickel-cadmium and lead-acid cells' output voltages continue to droop throughout their discharge. At discharge rates which expend the stored energy in approximately one hour, the silver-zinc cells provide from 30 to more than 50 watt-hours per pound, whereas the nickel-cadmium and lead-acid cells provide less than 10 watt-hours per pound.

The major limitation to the use of silver-zinc cells has been their comparatively short charged-storage life. Through continued research, the charged-storage life of secondary silver-zinc cells has been extended to more than a year. In applications where the batteries are not required to be kept in a constant state of readiness, battery life can be extended by refrigerated storage in the discharged condition.

Two by-products of the silver-zinc cell are of concern in their application. First, during charged storage, hydrogen is liberated at the cathode as a result of oxidation of the zinc by water. The amount of hydrogen liberated is not great, but caution must be taken in handling of batteries after long periods of charged storage. Second, near the end of high-rate discharges, the electrolyte boils quite violently and electrolyte is expelled from the cells. Thus installation provisions must be made to protect sensitive components from electrolyte spillage.

Paper 57-491, recommended by the AIEE Chemical Industry and Air Transportation Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 4, 1957; made available for printing May 23, 1957.

LYAL T. ABRAHAMSON is with Boeing Airplane Company, Seattle, Wash.

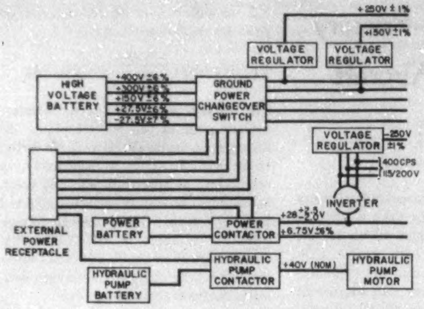


Fig. 1. Power-system block diagram

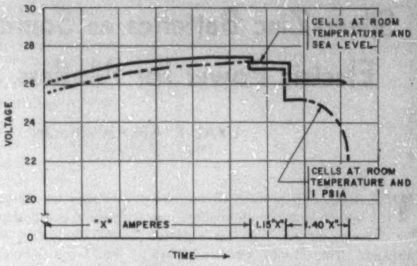


Fig. 3. Effect of altitude on cell voltage and capacity

Handling Techniques

Constant current charging consistent with the manufacturer's recommendations is used with all batteries. Whenever the cells are discharged for storage, rates taking from 2 to 4 hours are used to minimize heating of the cells.

Since the storage life of silver-zinc cells is limited, a constant search is made for handling techniques which will make the fullest use of their capabilities. The cells are purchased dry, in which condition they have been stored with no ill effects for over 3 years. When the need for a battery arises, the electrolyte is added and the formation cycle consisting of a low-rate charge and exhaustion discharge is performed. From this time until expenditure of the battery, deterioration continues. By storing the cells at approximately 20 to 30 F (degrees Fahrenheit) in the discharge condition, the useful life can be extended to three or four times what it is when the cells are stored at room temperature in the charged condition. Therefore the cells are maintained in the charged condition only when the need for a charged battery is present.

Another practical means found to increase the life of charged silver-zinc cells is replacing of the air in the battery cases with nitrogen. To investigate the benefits of nitrogen storage, 25 small cells (*HR-05*) were charged and placed in a sealed case containing nitrogen. After 3 months storage at 80°F, output efficiencies of from 67% to 81% were obtained at a 500-milliampere discharge rate. Recharge of the cells showed the input capacity was 81% of what it had been preceding the test. For comparison, 25 control cells were stored in air beside the nitrogen-filled case. At the end of the 3-month period, all the control cells failed to give any useful output as well as failing to respond to the recharge. Further tests using larger cells showed that the storage life could be increased more than 50% through the use of nitrogen-filled cases. Exclusion of oxygen from the cell apparently is the reason for the increase in cell life, since another series of tests using argon as a storage atmosphere gave similar results.

It has also been found that if the electrolyte is maintained at a high level in the cells, the life is increased markedly. However, electrolyte leakage then becomes a major problem and use of this scheme is impractical.

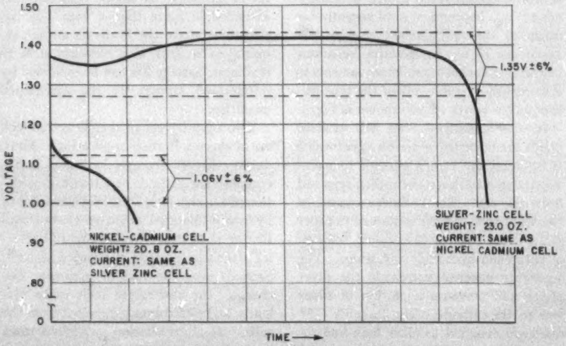


Fig. 2. Comparison of typical silver-zinc and nickle-cadmium cell discharge characteristics

Cell Selection

Two major considerations must be given in selection of a battery for any application. First, the battery must have sufficient ampere-hour capacity; and second, its voltage regulation must be consistent with the needs of the load.

The peak hydraulic power demanded is approximately three times that of the steady-state load. To keep the pump speed within the range of 6,200 to 10,000 rpm, the combined motor and battery weight was found to be lightest using a battery with excess ampere-hour capacity. In contrast, the loads of the other three batteries are more constant, and cells chosen for proper voltage regulation also have approximately the right ampere-hour capacity.

Application Problems

As with all components considered for installation in any aircraft, evaluation tests which include vibration, shock, temperature, humidity, altitude, and life tests are performed on the batteries. Two environmental conditions for which application modifications were found advisable are temperature and altitude.

During development of the battery for the +28-volt and 6.75-volt systems, protection of the electronic packages in ad-

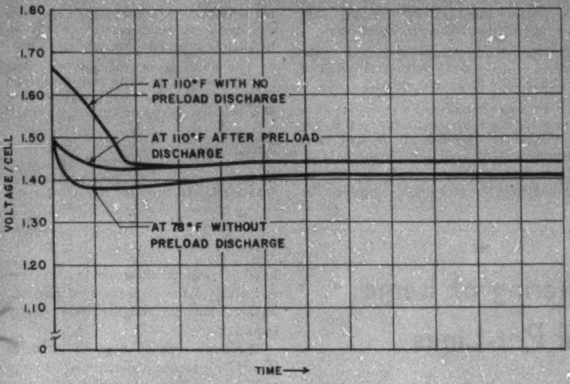


Fig. 4. Effect of temperature and preload discharge on cell voltage

jacent areas from contamination by electrolyte fumes was originally planned to be provided by completely sealing the battery case. The first test conducted under full-load conditions resulted in complete destruction of the battery. As the test progressed, the electrolyte reached its boiling point which in turn increased the internal case pressure and the electrolyte boiling point until the cell-case-material softening temperature was exceeded. An alternate scheme to exhaust the fumes overboard was then simulated in an altitude chamber. Both the voltage regulation and available ampere-hour capacity suffered severely at high altitudes. A third test finally was conducted to better define the effect of altitude. One group of cells in a bell jar was electrically connected in series to another group of cells outside the bell jar. After evacuation of the bell jar to less than 1 pound per square inch absolute, the design loads were applied and the individual cell voltages recorded. The results of the latter test are shown in Fig. 3. To provide both protection from compartment contamination by electrolyte fumes and to make full use of the battery's potentialities, a pressure-relief valve set for approximately sea-level pressure was placed in the exhaust line.

Two changes in the discharge characteristics of silver-zinc cells occur as the ambient temperature is varied. The plateau voltage increases approximately 1% for every rise of 20 F in temperature between 60 F and 120 F. In addition, and of more concern in maintenance of ±6% voltage tolerances, the shape of the discharge characteristic changes as the ambient temperature rises. Even at high-discharge rates, the higher voltage of the silver peroxide in the positive plates becomes predominate during the early part of the discharge if the ambient temperature is high. Preload discharging of the cells prior to use, if properly done, minimizes the dominance of the silver peroxide. Generally, expenditure of about 20% to 25% of the battery's capacity by the preload discharge is necessary to maintain less than ±6% voltage tolerance during the rest of the discharge. Fig. 4 shows discharges of cells at room temperature and 110 F without preload discharges, and one at 110 F after a preload discharge. To maintain voltages with the tolerances indicated in Fig. 1, the power and high-voltage batteries are given preload discharges before installation in the aircraft.

Battery Reliability

With lead-acid and other electrochemical systems, the battery's condition can be established by measurement of the short-circuit voltage or electrolyte specific gravity. These methods do not provide a satisfactory indication with silver-zinc cells since the internal resistance of silver-zinc cells is too low for a short-circuit test and its does not vary appreciably with the age of the cells. Also, the electrolyte is not available for test.

During the processing of thousands of cells, it has been found that failures do not occur until the input capacity decreases approximately 35% from what it is when the cells are new. Therefore, cell performance and reliability are maintained by using only cells which have an input capacity above a specified level. The cells are discharged and recharged at periodic intervals of less than 1 month for the small cells and up to 4 months for the large cells. The required frequency of these maintenance cycles has been found by measurement of the rate of energy loss during life tests conducted on each cell model. This method of cell maintenance has given 100% battery reliability to date.

Circuit Simplification

High-Voltage Battery Replacement

The high-voltage battery with its 320 individual cells is inherently an unreliable unit if age exceeds 60 days charged stand. To minimize the possibility of failures much painstaking maintenance is required. Although high-voltage battery reliability has been maintained at 100%, an active program for a replacement power supply was begun. First, the efforts were directed toward the use of dynamotors. Development problems of excessive broad-band noise and voltage drift could not be overcome, so the program was dropped.

With the advent of high-voltage high-power silicon diodes, development of a transformer-rectifier unit utilizing the 400-cycle-per-second inverter output became practicable. Not only is the high-voltage battery maintenance problem eliminated, but the system weight was reduced 15 pounds. The transformer-rectifier unit has the added advantage of enabling all final manufacturing testing to be done with the same high-voltage power source as that used in flight.

Alternate Power Sources

Periodic surveys and studies have been made of replacement systems for the secondary silver-zinc batteries. Included in these studies have been the ram-air turbine-driven alternator, the gas turbine-driven alternator and the primary silver-zinc battery systems. Until recently, the gas-turbine-driven alternator and conversion-unit development not only was uncertain but the development time and cost were prohibitive. With improvements in components it now appears that the development of the gas turbine and its power conversion units is feasible for future system designs.

The primary silver-zinc battery system which has been delayed by the lack of a positive automatic activation system also appears practical at this time. In many ways it is the best replacement for the present system, since it requires no maintenance. However, reliability is the major factor in selection of a replacement system and much statistical data will be needed to determine the reliability of the primary battery activation systems.

Conclusions

1. The silver-zinc battery system provides a power source for use in pilotless aircraft which is both light and reliable.

2. Battery life can be extended 50% through the use of nitrogen-filled cases.

3. Use of battery cases with pressure-relief valves overcomes the detrimental effect of altitude on voltage and capacity.

4. Preload discharges are required to maintain close voltage tolerances if the batteries are operated at elevated temperatures.

5. Reliable battery performance can be maintained by periodically putting the batteries through a discharge-charge cycle and discarding those batteries which fail to accept a pre-established minimum input.

6. Simplification of the power system and its maintenance is now possible through the use of a gas turbine-driven alternator or a primary silver-zinc battery system.

Reference

1. Primary Batteries (book), G. W. Vinal. John Wiley & Sons, Inc., New York, N. Y., 1950, p. 266.

Direct-Current Metering of Large Electrochemical Pot-Lines

M. E. REAGAN
FELLOW AIEE

Synopsis: D-c metering of large aluminum reduction pot-lines represents a major problem. It is no easy task to get the required degree of accuracy in measuring powers of 100,000-kw direct current. Taking advantage of the magnetic-core improvement in materials made available by developments during World War II, the magnetic amplifier emerged as a tool for many applications. Among these was a major improvement of the metering accuracies for large electrochemical pot-lines. In this paper the many advantages are discussed and some conclusions reached as to the best fields of application.

IN THE ENDEAVOR to obtain greater operating efficiency, engineers designing equipment for the production of aluminum are constantly seeking new methods and new components which may be beneficially applied. Playing a major role are efforts to obtain more accurate and reliable d-c metering. More accurate metering means more accurate control, which in turn leads to more aluminum production per kilowatt-hour.

The d-c power supply for the modern pot-line is from rectifiers where many individual units are paralleled to supply the total pot-line current. Fig. 1 illustrates one unit of a typical rectifier pot-line diagram. The largest one at present operates at 125,000 amperes and 800 d-c volts, being fed from 24 individual rectifiers averaging about 5,200 amperes each. The metering of total current on any line necessitates some method of adding the individual currents. Totalizing shunts of a size required to operate in the main totalizing bus are difficult to build and install.

The aluminum reduction plants built during World War II were normally rated at 50,000 to 60,000 amperes and 600 to 650 volts. Rectifiers were used with individual shunts which fed individual ammeters mounted on their individual control boards. The individual control boards were often built as a composite duplex control unit which necessitated long runs of the shunt leads and requiring special calibrations. Since these leads operated at full bus potential, the duplex board and the operators and maintenance personnel were subjected to the electrical hazards involved.

For the following purposes the current transducer was produced: 1. to totalize readily the pot-line current; 2. to facilitate running long distances to the individual ammeters; and 3. to insulate properly the switchboard instruments from the power conductors.[1] However, other magnetic-amplifier types of transductors were developed which were much smaller in physical size, much lower in energy loss, had greater accuracy and the ability to read zero, and which were less expensive. The relatively small magnetic amplifier also permitted a much greater ease of installation as it did not have to surround the main bus.[2,3] Due to quality of the magnetic-core material, its accuracy and stability were greatly improved.

The same type of magnetic amplifier is used for metering of potentials. A large value of resistance is connected in series with the control winding. Here also, the insulation which isolates the instruments from the power conductors is very beneficial. The readings are usually duplicated on the duplex control board and in the pot-room. Potential magnetic-amplifiers are not only used to meter the full pot-line voltage but also for portions of the pot-line to expedite the location of trouble in individual cells.

The combination of direct voltage and current metering quantities may be used to indicate and record kilowatts. Recording instruments for d-c kilowatts, volts, and total pot-line amperes are valuable in keeping the operating efficiency at a high level. Besides providing the permanent record, recording instruments often prove valuable by registering conditions immediately prior, during and immediately after troubles of various sorts have occurred.

Totalizing current metering is an important indication of over-all pot-line performance. Seven different methods have been used and others are being proposed. These do not vary radically from each other. In the large-sized stations, the use of shunts in the main bus is not practical at present.

Scheme 1. The first scheme consists of a large transductor mounted around the main d-c bus. The direct current is measured in terms of an alternating current which is within standard instrument range. The direct current in the main bus serves as the control element for saturating the magnetic circuit. When the direct current is zero, the ammeter would read a minimum value (exciting current) but not zero. Fig. 2 shows a schematic diagram of construction and a diagram of connections. This system should perhaps be used only where in-

---

Paper 57-228, recommended by the AIEE Chemical Industry Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 21–25, 1957. Manuscript submitted October 15, 1956; made available for printing February 15, 1957.

M. E. Reagan is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

The author wishes to acknowledge the co-operation and assistance of the engineering representatives of the Aluminum Company of America, the Anaconda Aluminum Company, the Kaiser Aluminum and Chemical Corporation, and the Reynolds Aluminum Company, who made this report possible.

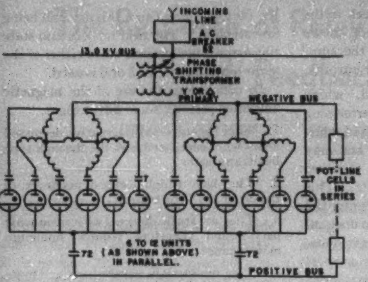

INCOMING LINE
A C BREAKER 52
13.8 KV BUS
PHASE SHIFTING TRANSFORMER
Y OR Δ PRIMARY
NEGATIVE BUS
POT-LINE CELLS IN SERIES
6 TO 12 UNITS (AS SHOWN ABOVE) IN PARALLEL.
72
72
7
7
POSITIVE BUS

Fig. 1 (left). Schematic diagram of rectifier group feeding pot-line

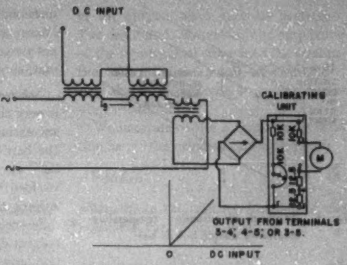

D C INPUT
CALIBRATING UNIT
10K
10K
10K
M
OUTPUT FROM TERMINALS 3-4; 4-5; OR 3-5.
O
DC INPUT

Fig. 3 (right). The input-output characteristic and schematic diagram of modern magnetic amplifier

dividual rectifier or generator outputs are not used, or as a check for a summation scheme of the individual rectifier outputs. The losses are comparatively large (the excitation for a 100,000-ampere unit being 110 amperes at 240 volts, single phase) and the first cost about ten times that of the magnetic amplifier in scheme 5. The published accuracy is 2% although field tests have shown it to perform within a 1% band of full-scale reading.

Scheme 2. The second scheme for totalizing consists of adding the outputs of a set of individual through-type transductors which measure the output of each rectifier or generator. The installed price for 12 rectifier units would be of the order of three to four times thatof scheme 1, but is necessary where individual rectifier outputs must be metered. The total excitation losses of 16 units is roughly one third of the single unit on the main d-c bus.

Scheme 3. The third totalizing scheme utilizes a d-c shunt in the cathode circuit of each rectifier (or in the positive lead of each d-c generator) which varies the control current of a magnetic amplifier with load. The output of the magnetic amplifier, in a similar fashion to scheme 2, operates the individual ammeter as well as combining with the other units to register total pot-line current; see Fig. 3. Some operators object to the use of a shunt for initiating the current indications, claiming that shunts are a source of error. This situation was thoroughly investigated and the objection was found to be true in many cases. However, it was found that the error was not due to the shunt materials or the method of manufacture but mainly to mechanical stresses caused by mounting methods. As manufactured, the shunts are very accurate, well within their guaranteed accuracy of 0.25%. When mounted in a bus, the strain on the manganin fins from the weight of the bus and the distortion resulting from bolt-tightening caused errors which have been reported as high as 20%. Another source of shunt error has been due to solder disintegration at the junction of the manganin and copper end castings. In spite of this, it is felt that the shunt is the best and most accurate means of measuring direct current when not distorted by the physical mounting. In some cases, the shunt should be supported at both ends and a flexible connector used at one end to minimize variations in output as the bus tends to move with temperature variations. Even though the shunt is changed from its factory calibration, this may be compensated for completely by changing the output of the magnetic amplifier to the correct value. There is at present no evidence that the shunt input-output ratio changes after installation.

The method of connecting the magnetic amplifier to the shunt must be studied. Due to the a-c ripple in a rectifier output, the variable field pattern induces a small voltage in the loop ordinarily resulting from the shunt lead configuration. To minimize this error, it is necessary to use flexible shunt leads and twist them as tightly as possible.

Scheme 4. The fourth method is identical to scheme 3, except that a millivolt drop is taken from the cathode or d-c bus instead of the shunt. On rectifiers, this scheme is fully as satisfactory providing the ambient temperature does not vary over too wide limits. This offers a very satisfactory scheme for those who object to the use of shunts.

Scheme 5. The fifth method uses a magnetic amplifier mounted near the main bus, using a multiple-wire tap, one on each bar of the bus. At least 6 feet of bus bar (12 feet is more desirable) must be spanned to get satisfactory results; see Fig. 4. The tapped section should not have any joints in it. The control wire (one per bar) should be placed physically close to the bus but insulated from it. One installation has the control wire embedded in a shallow saw-cut near the top center of each bus bar. The wire is insulated from the bar by varnish. This assures that the wire temperature closely follows that of the bus bar. Another installation ran the control wire between the bus bars halfway between the top and bottom of the bars. In this location, any

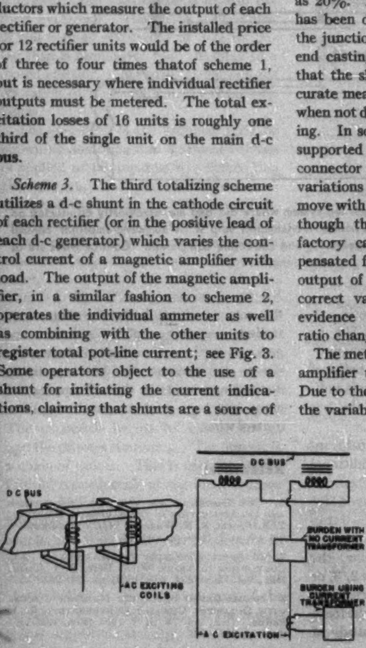

D C BUS
A C EXCITING COILS
DC BUS
BURDEN WITH NO CURRENT TRANSFORMER
BURDEN USING CURRENT TRANSFORMER
A C EXCITATION

Fig. 2 (left). Through-type transductor schematic diagram, and diagram of connections

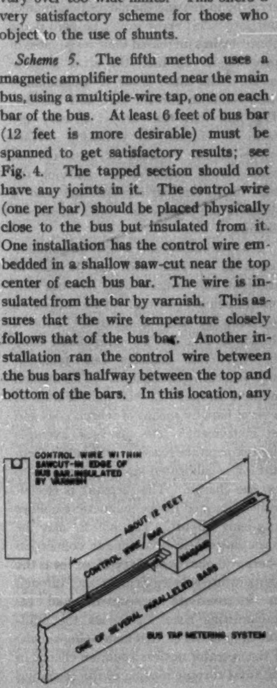

CONTROL WIRE WITHIN SAWCUT-IN EDGE OF BUS BAR.INSULATED BY VARNISH
ABOUT 12 FEET
CONTROL WIRE/BAR
MAGAMP
ONE OF SEVERAL PARALLELED BARS
BUS TAP METERING SYSTEM

Fig. 4 (right). One of several schemes for reducing temperature error in bus tap method of totalized current metering, by burying control wire in bus bar

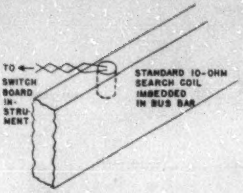

TO SWITCH BOARD IN-STRU-MENT
STANDARD 10-OHM SEARCH COIL IMBEDDED IN BUS BAR

Fig. 5. Method of mounting temperature-measuring device in bus bar

induced voltages in the control wires, by the small a-c component, is minimized.

Again referring to Fig. 3, the millivolt drop between the bus tap points will vary with the current in the bar. This millivolt drop will, with the millivolt drop of the other bus bars, determine the total magnetic-amplifier control current, and hence the magnetic-amplifier output. This is as it should be; however if the bus bar currents remain constant and the ambient temperature rises, the millivolt drop will increase, which in turn gives an error by increasing the magnetic-amplifier output. This error would be zero if the control wires increased their resistances in proportion. In fact, the error is quite small if the same grade of aluminum is used for both control wire and bus bar. The fact that the current density and ratio of current to radiating surface varies slightly between the wire and bar causes some small error. Burying the wire in the bus bar obviates the need of temperature connection under these conditions.

The output of the magnetic-amplifier tends to increase with an increase of ambient temperature with the bus current remaining constant. One suggestion for reducing this temperature error is shown in Fig. 5. To compensate fully for the error requires bus temperature measuring indicators. The manual method samples the bus temperature of several bars by means of search coils and an instrument. The operator averages the temperature readings and sets a rheostat on a calibrated scale to compensate for the change. The rheostat is electrically connected in the magnetic-amplifier totalizing circuit. The action may be made automatic by a similar connection where a 10-ohm exploring coil serves as the temperature-sensing element and, through an intermediate instrument and retransmitting rheostat, varies the millivolt drop in the current measuring circuit. If the operator notices a sudden change in the total current reading of the recording instrument, he may quickly check if the change was due to an "anode effect" in the pot-room or due to trouble in the summation of the individual rectifier outputs.

Scheme 6. A sixth method, which is very similar to scheme 5, is performing successfully; see Fig. 6. It is used when the bars of the main totalizing bus may be electrically connected together (welded) at the two tap points. This assures that there can be no potential difference between the bars due to unequal current distribution. When this is done, one large control wire (properly proportioned in size) may be used. This scheme also has the advantage of quickly recognizing an open circuit in the control wire. When multiple-control wires are used, one open circuit may temporarily be mistaken for an "anode" effect.

Scheme 7. A seventh method has been tried (in Europe) and may show some promise. It has been developed for d-c magnitudes of some 15,000 amperes. It is known as the Hall generator. The d-c bus is surrounded by an iron magnetic circuit with air gaps. Hall generators are located in the air gaps and their voltage outputs are proportional to the bus current.

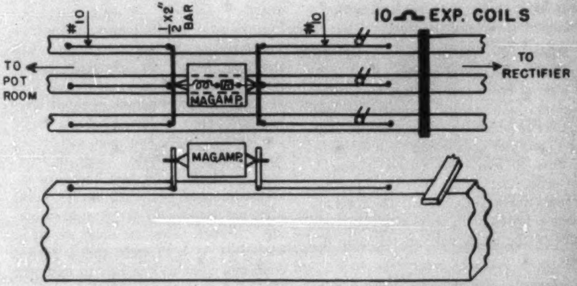


Fig. 6. Scheme of connections for bus tap system where bus bars are electrically connected on source side between tap points and source. Manganin strip R in magnetic-amplifier control winding requires use of temperature measuring and correcting equipment. Control wires are of equal length

Operating reports from the field installations indicate that an additional error appears under conditions when one or two rectifier units are removed for maintenance purposes. If, by tap changer adjustment, the direct voltage is brought back to its original value, the summation of current by methods 2, 3, or 4 should also be expected to return to its former level. The field tests indicate a lower current in all cases. The study of the factors which cause this are still being investigated. The field reports also state this error is not present when either of the main bus tap systems 5 or 6 is used.

In conclusion, the use of the magnetic amplifier permits some advantages:

1. They operate with greatly decreased losses when compared with the through-type bus transductor.
2. They have a high degree of accuracy, of the order of 1% or better.
3. They are static devices (with no moving parts) and therefore require minimum maintenance.
4. They are small in physical size and more easily mounted.
5. They isolate the switchboard instruments from the high-voltage d-c power busses.
6. Their use permits extremely long physical distances over which the current indications may be transmitted.

Where individual currents, as well as the totalized d-c bus current, is needed, the scheme of totalizing the shunt-controlled magnetic-amplifier outputs appears to be the second best arrangement. Up to the present, the best totalizing method appears to be that of tapping the main bus bars with magnetic-amplifier control wires.

References

1. The Theory of the Current Transductor and Its Application in the Aluminum Industry, T. R. Specht, R. N. Wagner. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 441–52.

2. Magnetic-Amplifier Applications in D-C Conversion Stations, W. A. Derr, E. J. Cham. *Ibid.*, vol. 72, pt. III, Apr. 1953, pp. 220–29.

3. Single-Ended Saturable Reactor Circuit with Quiescent Current Compensation, R. J. Radus. *Ibid.*, vol. 73, pt. I, 1954 (Feb. 1955 section), pp. 597–603.

The Response of Relay Amplifiers with Feedback

JOHN E. GIBSON
ASSOCIATE MEMBER AIEE

FRANZ B. TUTEUR
ASSOCIATE MEMBER AIEE

THE SIMPLICITY and economy of on-off controllers has led to their wide control system application in industry. It is possible to employ the on-off controller principle in the design of a power amplifier that is simple and rugged. The use of such a device will be restricted to the operation of loads that incorporate a smoothing or low-pass effect in their response.

Basic Operation

The basic configuration of the relay amplifier is shown in Fig. 1. The relay is a polarized relay or one that closes to one contact for a positive coil voltage and to a second contact for a negative-coil voltage. A deadband or region in which neither contact is closed is typical; and usually the relay will exhibit hysteresis. That is, a larger value of coil voltage is required to close the relay than is required to hold the relay in a closed position. The filter is required for operation of the amplifier as will be seen below.

The operation of the amplifier may be shown by assuming that a step of voltage is applied to the input at t_0. Before this time the relay was open and there was no charge on the capacitor. The input voltage begins to charge the capacitor and as shown in Fig. 2, at t_1 the voltage e_r has reached a large enough value to close the relay. At t_1 then, e_m suddenly jumps from 0 to E_b. That is, the relay contact closes and the output is connected to the positive supply. The capacitor must now begin to charge to $e_{in}-E_b$ and a downward exponential begins. At t_2 the capacitor voltage has become so small that the relay cannot hold and the contact falls open. The capacitor begins to charge to e_{in} and the process continues. The output is a chain of pulses. This is certainly not a faithful reproduction of the input-voltage wave. However, if the load acts like a low-pass filter, such as a d-c motor for example, and is insensitive to frequencies as high as the pulse repetition rate, the amplifier is satisfactory and in fact, as shown below, is "linear" over most of the amplitude operating range. If the input-signal amplitude is large, the pulses are long and the average value of the output is high. The upper limit on input-signal amplitude is reached at the point at which the relay remains closed for the entire period. This limit is determined by the value of the supply voltages on the relay contacts. The lower limit on signal amplitude is determined by the value of the relay deadband.

Relay Amplifier Stability

Before discussing the input-output relations of a relay amplifier, the stability of such a device will be considered. Relays have been used directly in closed-loop systems and the design and stability criteria for relay servos have been discussed by Kochenburger in terms of a describing-function analysis.[1] The analysis of a relay amplifier may be carried on using these describing-function techniques. In Fig. 3(A) is shown a block diagram of a relay amplifier. The block H_1 represents a linear electronic amplifier and the block H_2 represents the relay, which is nonlinear. The relay static characteristic is given in Fig. 3(B). For the type of resistance adder circuit shown, a reversal of sign in H_1 or H_2 is necessary to give negative feedback.

The relay is assumed to have deadband (Δ) and hysteresis (h) as defined in Fig. 3(B). The Kochenburger describing function may be used to obtain the quasi-linear transfer function of the relay from the magnitude of a sinusoidal driving signal to the magnitude and phase angle of the fundamental of the output wave,

$$H_2=\frac{E_{m1}}{i_r}=\frac{4\sin\beta}{\pi I_r}\underline{/-\alpha} \tag{1}$$

where

$$i_r=I_r\cos\omega t \tag{2}$$

and

$$\begin{aligned}\alpha&=\frac{1}{2}\left\{\cos^{-1}\frac{\Delta-h}{2I_r}-\cos^{-1}\frac{\Delta+h}{2I_r}\right\}\\ \beta&=\frac{1}{2}\left\{\cos^{-1}\frac{\Delta-h}{2I_r}+\cos^{-1}\frac{\Delta+h}{2I_r}\right\}\end{aligned} \tag{3}$$

The derivations of equations 1 and 3 are given by Kochenburger and require only the assumption that the higher harmonics of the output wave may be neglected. The deadband, hysteresis, and amplitude of the input signal enter into the evaluation of both α and β and, as seen below, may shape the response so as to contribute to instability.

We may also write a quasi-linear relation for A, the negative of the loop gain

$$\hat{A}=\hat{H}_1\hat{H}_2\hat{H}_3 \tag{4}$$

where H_3 is the transfer function of the resistance-capacitance network

$$H_3=\frac{R_1}{(R_1+R_2)\left[1+\frac{R_1R_2CS}{R_1+R_2}\right]} \tag{5}$$

The loop is stable if the roots of the characteristic equation $(1+A)$ lie in the left half of the s plane. Let us assume that $R_1=R_2$ and that H_1, the amplifier, is perfect and has a gain of K. Then,

$$H_1H_3=\frac{KR_1}{2R_1(1+R_1CS/2)}=\frac{K'}{1+TS} \tag{6}$$

Since

$$\frac{y}{x}=\frac{H_1H_2H_3}{1+H_1H_2H_3} \tag{7}$$

we may write

$$\frac{y}{x}=\frac{H_2}{\frac{1}{H_1H_3}+H_2} \tag{8}$$

According to the Routh or Leonard stability criteria, the plot of the denominator must progress around the origin keeping the origin to its left for stability as shown in Fig. 4(A). However, since H_2 depends on the amplitude of the input signal, it is simpler to plot $-H_2$ and $1/H_1H_3$ starting at the origin as in Fig. 4(B). In this figure stability is indicated by the fact that the two curves do not intersect. The frequency sensitive portion $1/H_1H_3$ is the inverse Nyquist diagram. The plot may be inverted as shown in Fig. 4(C). This is superior because the important portion of the amplitude function is expanded and the accuracy of the plot in the neighborhood of an intersection is improved. In any case the system shown in Fig. 4 is stable because there is no intersection of the loci.

In Fig. 5 is shown a specific plot for a relay amplifier with $\Delta=0.15$, $h=0.05$, and $I_{r\max}=1$. It would be assumed from

---

Paper 57-183, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N.Y., January 21-25, 1957. Manuscript submitted August 24, 1956; made available for printing February 15, 1957.

JOHN E. GIBSON and FRANZ B. TUTEUR are with Yale University, New Haven, Conn.

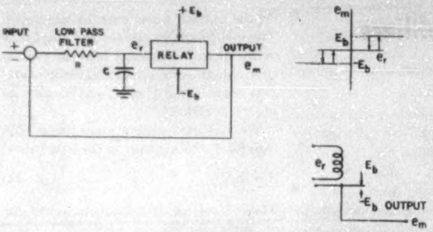


Fig. 1 (above). The basic configuration of the relay amplifier

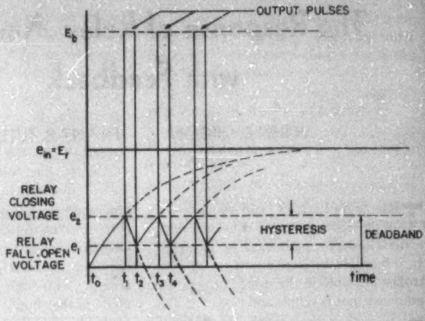


Fig. 2 (right). Voltage wave shapes in response to a step function input

the figure that under no condition of gain K or amplitude of input signal could this system be unstable. However at high gains, previously neglected lags in the amplifier or in the relay itself may cause the frequency locus to bend around as shown in Fig. 4(B), thus possibly causing instability. Fig. 5 also shows, indirectly, the effect of increasing the break frequency of the filter. If the filter-break frequency is raised to a value such that these neglected lags in the loop become important, the frequency locus will bend to the left around the origin. If it intersects the amplitude locus, instability is indicated.

Relay Amplifier Static Characteristic

The output of a relay amplifier with feedback is a series of pulses, positive or negative, of amplitude E_b and with length and repetition rate determined by the input. The output wave shape is completely independent of the load, and nowhere in the following analysis need the load be specified. Usually, however, the relay amplifier is employed in an application where only the average value of this train of pulses is important.

In order to determine the static relation between the input voltage e_r and the average output voltage E_m, we will consider H_1 and H_2, in Fig. 3, as grouped together. If the amplifier H_1 is linear, the static characteristic between e and e_m may be described as in Fig. 6. Gain in H_1 reduces directly the effective values of deadband and hysteresis, e_1 and e_2.

Let us suppose at zero time, t_0, that $e=0$. Thus the relay is open and $e_m=0$. Let a step voltage, E_r in magnitude be applied as shown in Fig. 2. Then

$$e = E_r(1 - \epsilon^{-t/T_1}) \tag{9}$$

where $T_1 = R_1 C$. When e reaches e_2, the relay closes. Substituting e_2 for e in equation 9, we may solve for t_1, the time of closing.

$$\epsilon^{-t_1/T_1} = \frac{E_r - e_2}{E_r} \tag{10}$$

or

$$t_1 = -T_1 \ln(1 - e_2/E_r) \tag{11}$$

After the relay closes,

$$e = E_r\left(\frac{R_2}{R_1+R_2}\right)[1 - \epsilon^{-(t-t_1)/T_2}] - E_b\left(\frac{R_1}{R_1+R_2}\right)[1 - \epsilon^{-(t-t_1)/T_2}] + e_2\epsilon^{-(t-t_1)/T_2} \tag{12}$$

where $T_2 = (R_1R_2/R_1+R_2)C$. When e reaches e_1, the relay opens. The time of opening, t_2, may be found by setting $e=e_1$ and solving for $t=t_2$,

$$\epsilon^{-(t_2-t_1)/T_2} = \frac{e_1 - E_r\left(\frac{R_2}{R_1+R_2}\right) + E_b\left(\frac{R_1}{R_1+R_2}\right)}{e_2 - E_R\left(\frac{R_2}{R_1+R_2}\right) + E_b\left(\frac{R_1}{R_1+R_2}\right)} \tag{13}$$

thus

$$t_2 - t_1 = -T_2 \times \ln\left[\frac{e_1 - E_r\left(\frac{R_2}{R_1+R_2}\right) + E_b\left(\frac{R_1}{R_1+R_2}\right)}{e_2 - E_r\left(\frac{R_2}{R_1+R_2}\right) + E_b\left(\frac{R_1}{R_1+R_2}\right)}\right] \tag{14}$$

Equation 15, following, is obtained when the relay is open:

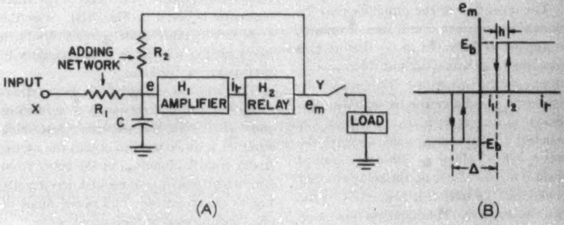


Fig. 3. A typical amplifier: (A) block diagram, (B) relay characteristics

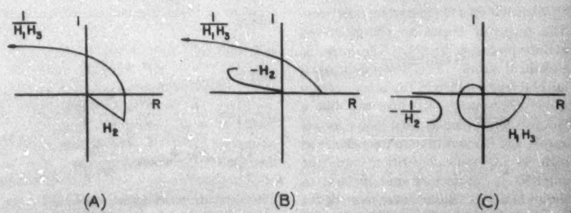


Fig. 4. Stability plots for a typical relay servo. The three plots show the same system

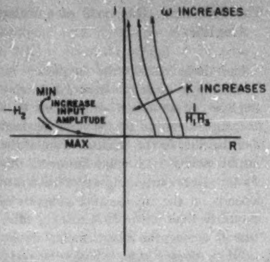


Fig. 5. Stability plot for an actual relay amplifier showing the effect of gain variation

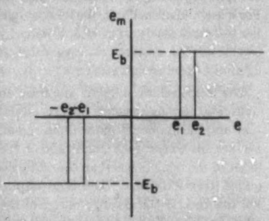


Fig. 6. Characteristic of linear amplifier and relay, deadband expressed in terms of error voltage

$$e=E_r(1-\epsilon^{-(t-t_2)/T_1})+e_1\epsilon^{-(t-t_2)/T_1} \qquad (15)$$

It will be noted that equation 15 is the same as equation 9 plus the term due to the initial voltage on the capacitor. All of the segments of the voltage curve will be exactly the same as these described by equations 12 and 15. Equation 9 is the special case of no charge on the capacitor. The relay closes again at t_3 when $e=e_2$ given by

$$t_3-t_2=-T_1\ln\left(\frac{e_2-E_r}{e_1-E_r}\right) \qquad (16)$$

During the period of time (t_3-t_2) the output voltage is zero and during the period (t_2-t_1) the output is $-E_b$. These two lengths of time represent the cycle of the output-voltage wave. The average output is therefore

$$\overline{e_m}=\frac{-E_b(t_2-t_1)}{t_3-t_1} \qquad (17)$$

Substituting equations 14 and 16 into equation 17,

$$\overline{e_m}=\frac{E_b\ln\dfrac{e_1-E_r\left(\dfrac{R_2}{R_1+R_2}\right)+E_b\left(\dfrac{R_1}{R_1+R_2}\right)}{e_2-E_r\left(\dfrac{R_2}{R_1+R_2}\right)+E_b\left(\dfrac{R_1}{R_1+R_2}\right)}}{\dfrac{T_1}{T_2}\ln\dfrac{e_2-E_r}{e_1-E_r}+\ln\dfrac{e_1-E_r\left(\dfrac{R_2}{R_1+R_2}\right)+E_b\left(\dfrac{R_1}{R_1+R_2}\right)}{e_2-E_r\left(\dfrac{R_2}{R_1+R_2}\right)+E_b\left(\dfrac{R_1}{R_1+R_2}\right)}} \qquad (18)$$

By a relatively minor change in the circuit, R_2 may be grounded whenever the relay is open. This will make $T_1=T_2$ and will contribute to the linearity of the input-output relation. In this case equation 18 reduces to the following:

$$\overline{e_m}=\frac{E_b\ln\left[\dfrac{1-\left(\dfrac{2e_1}{E_r-E_b}\right)}{1-\left(\dfrac{2e_2}{E_r-E_b}\right)}\right]}{\ln\left[\dfrac{1+\dfrac{1}{2}\left(\dfrac{E_b+E_R}{e_1-E_R}\right)}{1+\dfrac{1}{2}\left(\dfrac{E_b+E_R}{e_2-E_R}\right)}\right]} \qquad (19)$$

In certain applications an electronic difference amplifier is used rather than the resistor adding network. The low-pass filter may be incorporated as a part of the interstage coupling network of the electronic amplifier. When the difference amplifier rather than the resistor adding network is employed, the factors representing voltage divider action do not appear and equation 18 may be simplified to yield

$$\overline{e_m}=\frac{E_b\ln\left[\dfrac{1-\dfrac{e_1}{E_r-E_b}}{1-\dfrac{e_2}{E_r-E_b}}\right]}{\ln\left[\dfrac{1-\dfrac{E_b}{E_r-e_1}}{1-\dfrac{E_b}{E_r-e_2}}\right]} \qquad (20)$$

By examination of this relation we can see that it gives the negative real result that is required, only for the region, $e_2\le E_r\le E_b+e_1$. For $E_r<e_2$ the relay does not close: the output is zero. For $E_r>E_b+e_1$, the relay remains closed and the output is E_b. It may be shown that if e_1 and e_2 are

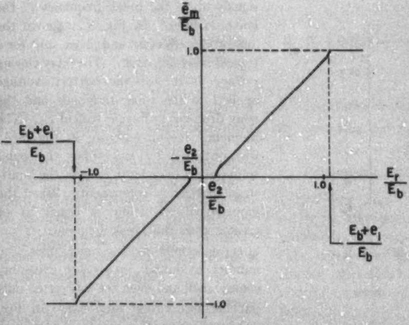


Fig. 7 (above). Static input-output characteristic of a relay amplifier

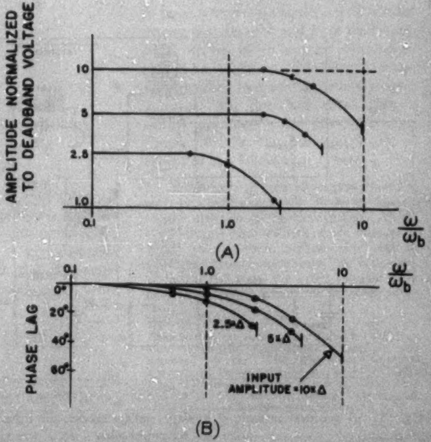


Fig. 8 (right). Frequency response of relay amplifier: (A) amplitude normalized to the value of deadband versus angular frequency to a log-log-scale, (B) phase lag versus angular frequency to a semi-log scale small with respect to the battery voltage, the slope of e_m versus E_r is essentially constant over the operating range. The exact relation between input voltage magnitude and the average of the output voltage may be determined by the use of equation 20. In Fig. 7 is shown $\overline{e_m}$ versus E_r normalized to E_b.

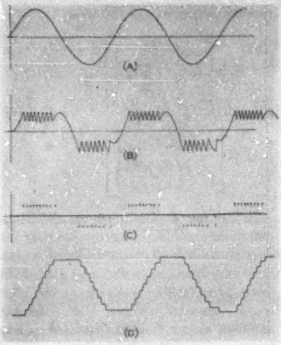


Fig. 9 (left). Voltage wave shapes in response to sinusoidal driving signal: (A) sinusoidal input signal, the frequency is the same as the break frequency (ω_b) of the low-pass filter; (B) output of low-pass filter; (C) output of relay amplifier; (D) integrated output of relay amplifier

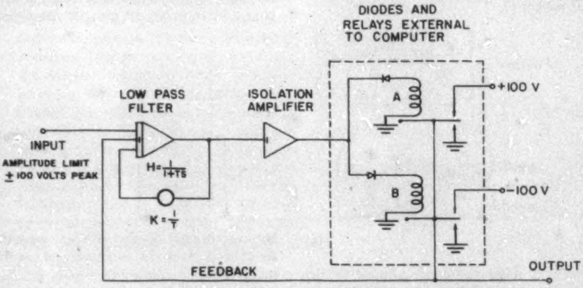


Fig. 10. Circuit for obtaining frequency response and wave shapes

Relay A: Deadband, 10.07 volts
Hysteresis, 5.17 volts

Relay B: Deadband, 9.9 volts
Hysteresis, 5.2 volts

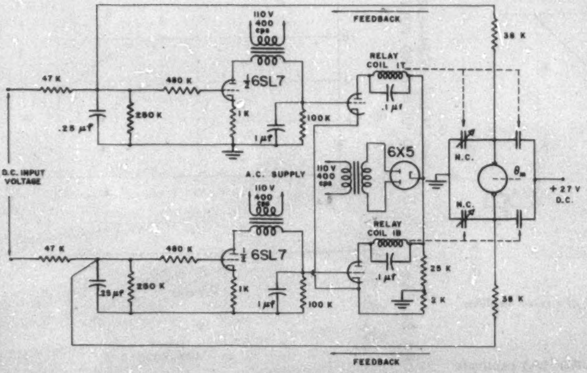


Fig. 11. A practical example of a relay amplifier used to drive the positioning motor of a servomechanism

The Frequency Response of a Relay Amplifier

The limitation on the frequency response, for a sinusoidal input, of a relay amplifier is essentially determined to a first approximation by the time constant T of the filter or the "chatter" rate of the output pulses. Actually, however, this chatter rate or output pulse repetition rate depends on the input-signal magnitude, and in the limit where E_r is either greater than E_b or zero the relay remains always closed or always open. Since the relay amplifier is a nonlinear device, the entire concept of frequency response is contrived, and a measure for one type of input signal will not hold for another. For a step function input, the relay amplifier is particularly fast and essentially establishes the required average value of $\overline{e_m}$ in one opening and closing.

The response of a relay amplifier to sinusoidal signals depends on the amplitude and the frequency of the input signal. For signals whose peak amplitude at the input of the relay is less than the value required to close the relay, there is no output. For instance, if the closing voltage of the relay in 10 volts, a signal whose peak amplitude is less than 10 volts will not close the relay. Since the low-pass filter attenuates the amplitude of the input signal, the peak amplitude must be computed at the input to the relay. At the break frequency of the filter, for instance, the filter reduces the input amplitude by a factor of 0.707, thus the minimum input-signal amplitude must be 14.14 volts peak. Essentially a relay amplifier will operate properly for normal amplitude signals up to the break frequency of the low-pass filter. In Fig. 8 is shown the amplitude reduction and phase shift for a typical relay amplifier. The relay closing voltage is 0.1 of the battery voltage applied to the relay terminals and the relay drop-out voltage is 0.05 of E_b. The frequency response shown in Fig. 8(A) was obtained by smoothing the pulse train output in order to recover the input-signal frequency component. When the input-signal frequency is equal to or greater than the break frequency of the filter, especially at low amplitudes, the number of output pulses per cycle becomes small and even the integrated output becomes rather step like. In Fig. 9 the various wave shapes for a particular signal are shown. In 9(A) is shown the sine-wave input to the amplifier. Its peak amplitude is 0.5 and its frequency is the same as the break frequency of the amplifier low-pass filter. In 9(B) is shown the wave form at the output of the low-pass filter which illustrates the deadband and hysteresis of the relay. In 9(C) is given the output of the relay, and in 9(D) is shown the integrated output. In a typical power-amplifier application the relay output drives a d-c motor which would have an integration plus a low-pass filter. Thus it could be expected that the wave shape in 9(D) would be smoothed by the filter before appearing at the output shaft.

The frequency response shown in Fig. 8 and the wave shapes shown in Fig. 9 were obtained with the circuit shown in Fig. 10. The relays used were single-pole double-throw and connected with diodes to form the polarized characteristic. The low-pass filter and the subtractor were synthesized on an analog computer, which made data taking and recording of wave shapes convenient. A practical example of a relay amplifier used to drive the actuating motors of an automatic field plotter[2] is shown in Fig. 11. Here separate amplifiers are used to drive the two relay coils required to complete the motor contacts.

It is difficult to establish an equivalent loop-gain constant for the relay amplifier with feedback without making several rather arbitrary assumptions. However, from Fig. 8 it can be seen that such an equivalent gain depends on the input-signal amplitude. At maximum amplitude signal the break frequency of the closed-loop response of the experimental amplifier occurs at about $4/T$ and for minimum amplitude signals the break approaches $1/T$. This would indicate a loop-gain constant of about 3 at maximum signal amplitude and zero for minimum amplitude signals.

The amount of phase shift contributed by the filter to the over-all phase shift shown in Fig. 8(B) is inversely related to the gain of the amplifier, due to the negative feedback. In addition to decreasing the closed-loop phase shift, an increase in amplifier gain will act to reduce the effective deadband, thus increasing the usefulness of the amplifier at low-signal amplitudes. As the gain is increased, the loop becomes less stable due to the increased effect of previously neglected lags in the amplifier and relay itself, until finally the amplifier goes into stable oscillation. That is, its output for no input is a continuous string of alternate positive and negative pulses of a predictable frequency. As a signal is placed on the input terminals, the length of the positive pulses grows longer and the negative pulses shorter or vice versa, thus the averaged output follows the input. This type of operation is more sensitive to low-amplitude signals than stable operation and will have minimum phase shift from input to output. However, the continuous vibration of the power element under quiescent conditions may be undesirable.

Conclusions

The static response to step function inputs to a relay amplifier with feedback has been shown to be linear over the normal operating range. The dynamic response of such amplifier has been discussed and the response to sine-wave inputs has been determined. Extension of these results to arbitrary inputs must be done with considerable care. The effects of raising the gain in the forward path of the amplifier have been discussed and the effect of relay deadband and hysteresis on the response of the amplifier has been shown.

References

1. A FREQUENCY RESPONSE METHOD FOR ANALYZING AND SYNTHESIZING CONTACTOR SERVOMECHANISMS, **Ralph J. Kochenburger.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

2. AN AUTOMATIC FIELD PLOTTER, **R. Gelfand, B. J. Shinn, F. B. Tuteur.** *Ibid.*, vol. 74, pt. I, Mar. 1955, pp. 73–78.

Heat Pump and Heating Cables Installed in the Same Residence for Data Purposes

C. W. JONES
MEMBER AIEE

E. E. LINDEN
MEMBER AIEE

THIS paper deals with an installation of electric heating and air conditioning in a New England residence. The installation was made through the close co-operation of the local utility and a family building a new home, and serves the particular purpose of providing the utility with information. It is probably unique because the one house features two full-capacity systems.

1. A water-to-air heat pump with supplementary resistance heaters.
2. Heating cables buried in the ceiling plaster.

The reason for this double system is to provide information on each separately, at the same time forming a close comparison of the two in terms of the following:

1. Installation features and costs.
2. Operating and maintenance costs.
3. Effect on the utility system.
4. Promotional features.

Realizing that many utilities are active in this field, as witnessed by a recent count of some 300,000 installations, this paper is offered as a contribution toward the efforts of the newcomers and for criticism by those already in the field.

Basic Approach

Careful thought was given to the implications in a pilot installation of this nature before planning the details. The utility involved had sponsored two experimental heat-pump installations 10 years ago and kept up to date on the subject even though the experimental heat pumps had not proved entirely suitable. With this background a list of basic notions was made, and it seemed proper that these should form the framework within which a pilot installation should be consistent. No effort is made to generalize beyond this utility's area; however, this does not mean that a similar summary would not apply elsewhere. Briefly the basic approach is shown in the following:

1. Public acceptance of electric heating and air conditioning will in large depend on the attitude of the local utility. Since electricity is the utility's most important if not sole product, perhaps any sort of negative attitude on this new load is unwise. If this is the case, the utility should then become as fully informed on the subject as

---

Paper 57-539, recommended by the AIEE Domestic and Commercial Applications Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE North Eastern District Meeting, Pittsfield, Mass., May 1–4, 1957. Manuscript submitted February 19, 1957; made available for printing May 23, 1957.

C. W. JONES and E. E. LINDEN are with the Narragansett Electric Company, Providence, R. I.

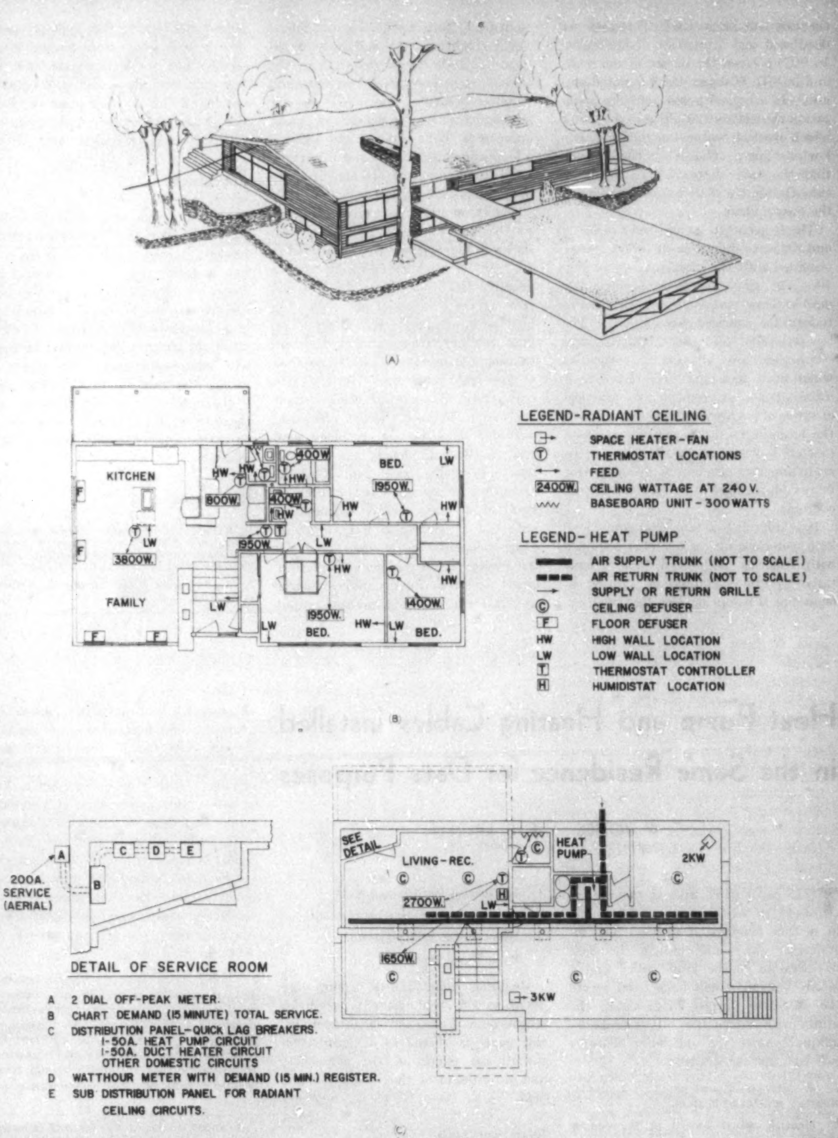

(A)
KITCHEN
FAMILY
BED.
BED.
BED.
800W.
400W
400W
1950W.
1950W.
1950W.
1400W.
3800W.
HW
LW
F
(B)
LEGEND-RADIANT CEILING
SPACE HEATER-FAN
THERMOSTAT LOCATIONS
FEED
2400W. CEILING WATTAGE AT 240 V.
BASEBOARD UNIT - 300 WATTS
LEGEND- HEAT PUMP
AIR SUPPLY TRUNK (NOT TO SCALE)
AIR RETURN TRUNK (NOT TO SCALE)
SUPPLY OR RETURN GRILLE
C CEILING DEFUSER
F FLOOR DEFUSER
HW HIGH WALL LOCATION
LW LOW WALL LOCATION
T THERMOSTAT CONTROLLER
H HUMIDISTAT LOCATION
200A. SERVICE (AERIAL)
A
B
C
D
E
DETAIL OF SERVICE ROOM
A 2 DIAL OFF-PEAK METER.
B CHART DEMAND (15 MINUTE) TOTAL SERVICE.
C DISTRIBUTION PANEL-QUICK LAG BREAKERS.
1-50A. HEAT PUMP CIRCUIT
1-50A. DUCT HEATER CIRCUIT
OTHER DOMESTIC CIRCUITS
D WATTHOUR METER WITH DEMAND (15 MIN.) REGISTER.
E SUB DISTRIBUTION PANEL FOR RADIANT CEILING CIRCUITS.
SEE DETAIL
LIVING-REC.
HEAT PUMP
2KW
2700W.
1650W.
3KW
(C)

Fig. 1. Plan of new home

A—Architect's sketch
B—First-floor plan
C—Basement plan

possible and later decide whether actively to promote or merely deal with inquiries if and when they arise.

2. Any utility policy on this subject should develop only after the results of enough local experience, so that such policy will have a valid basis. This might involve one or several pilot installations.

3. Automatic control should be fully developed in any pilot plant, whether it involves heat-pump or direct-conversion equipment. Packaged equipment should be preferred to specially fabricated equipment, the latter having little or no long-range implication.

4. Water-source heat pumps are simpler and more efficient over-all than the air-source heat pump; however, the application of water-source machines is generally limited to locations where a source of water is readily available and disposable. In suburban developments, where this is generally not the case, the air-source machine has its best application. It seems apparent that this gives the air-source machine a greater potential in numbers.

5. Direct-conversion electric heaters offer the simplest means of heating with complete zoning and low first cost; and, where there is no particular desire for cooling or dehumidification, acceptance will depend upon whether a home-owner will bear the cost of operation for the convenience, released space, etc. For supplementary heating, direct conversion has good potential.

6. The best application of either heat pump or direct conversion is in new construction, probably occurring first in homes priced $20,000 and higher. A pilot installation then should be in this price class, with both heat-pump and direct-conversion equipment, if possible, as well as other features which would be compatible with promotional use later on.

General Features of Site and Structure

Considerable time was spent in planning the details involved with the location and building features. Every effort was made to closely approximate the general features of site and construction which might be representative of the first few all-electrical homes in the area. The completion of this project allows for close study of those provisions which the utility must make in order to satisfactorily serve a load which includes full-appliance application and modern lighting as well as heating and air conditioning.

An attractive location suggests use for promotional purposes later on, however, it certainly does not detract from the all-electrical concept. The pilot installation should have an acceptable appearance to the average person and not look like a laboratory. There were favorable reactions from the various subcontractors involved. The newspaper ran a news story shortly after the structure was framed in, and later a home-section feature upon completion. As a result, the project has earned acceptance as the forerunner of things to come rather than experimentation.

Building Site, Brook, and Well

The selection of a building site was made after it was determined that a water-source heat pump would be used. A 4-acre piece of land, several miles beyond the suburbs was selected. It is well wooded with a somewhat sloped terrain and a brook. The brook is particularly attractive since it serves as a means of disposing of the water from the heat pump. Trees, for the most part, have been left to shade the house, carport, and parking area in the summer time. In winter, ample sunshine is available through the bare branches, serving to reduce the heating load on fair days.

An 8-inch well, 175 feet deep, presents a flow of 20 gpm (gallons per minute) without lowering its level beyond 65 feet. This is an unusually good well for the area, other similar wells delivering 5 to 10 gpm. Water temperature in the late winter is about 49 F (degrees Fahrenheit) and in the late summer about 51 F.

Architecture

In the architecture of the house, the resulting design would probably be termed functional-contemporary. Essentially, it is a ranch house set cornerwise into a slope. One end is for sleeping (upstairs) and storage (downstairs), and the other end (upstairs and downstairs) for living. Fig. 1(A) is a sketch of the house and Figs. 1(B) and (C) show the first-floor and basement plans. Glass is used sparingly on the bedroom end and functionally on the living-room end. The delightful feeling of openness that glass affords must be evaluated to some degree in terms of heat loss on a dark day when the wind is blowing at 30 to 40 miles per hour and the temperature is between 0 and −10 F. These are not accepted design conditions; however, such are of record. In this connection, the application of draperies which may be opened or closed in several sections reduces the heat losses through extensive glass areas as necessary, while still passing light.

The placement of living room in the basement is an attempt more fully to utilize the volume of the building for living. It is not an innovation. However, the application of a heat pump and heating cables to this type of area is certainly worthy of study.

Other Features

Insulation. The most important feature of the entire structure is the insulation. This not only reduces the heat loss and heat gain, but introduces thermal inertia which particularly minimizes the rate at which the structure cools off on cold days between heating cycles. This of course correlates to a high degree of comfort.

The economic justification of thermal insulation is well known. In this application, batt-type insulation with a vapor barrier was installed. There are 6 inches above the ceiling, 3 inches in the side walls, and 2 inches between normally heated and unheated spaces. This is essentially in accord with standards subscribed to by many manufacturers of electric heating equipment.

Windows are all double-glazed and weatherstripped. Doors to the outside are weatherstripped; storm doors have been added. Windows of the awning type in all rooms can be opened for natural ventilation on the nice days, as well as for reasons of safety.

Adequate Wiring. Wiring throughout is in accord with adequate wiring standards as a minimum.

Light Conditioning. Valence lighting when it is combined with a few functional fixtures creates an unusually pleasing effect.

Appliance Application. A full complement of electric appliances are in service.

Sizing the Heating and Air-Conditioning Equipment

Heat Loss and Gain

To size the equipment properly, heat loss and gain calculations were made by several engineers. The resulting figures differed widely on heating from 42,000 to 75,000 Btu per hour, and on cooling the difference was between 33,000 and 36,000 Btu per hour. An investigation of methods used to arrive at these figures explained the differences. A high factor of safety, a large number of air changes per hour, lack of faith in insulation values (based rightly or wrongly on the suspicion that equivalent foil insulation might be used), shading factors, and questionable design temperatures account in large for the high values. The low values are based on an almost opposite treatment. In any event this experience indicates the significance of the various judgment factors that enter into the calculation of gains and losses.

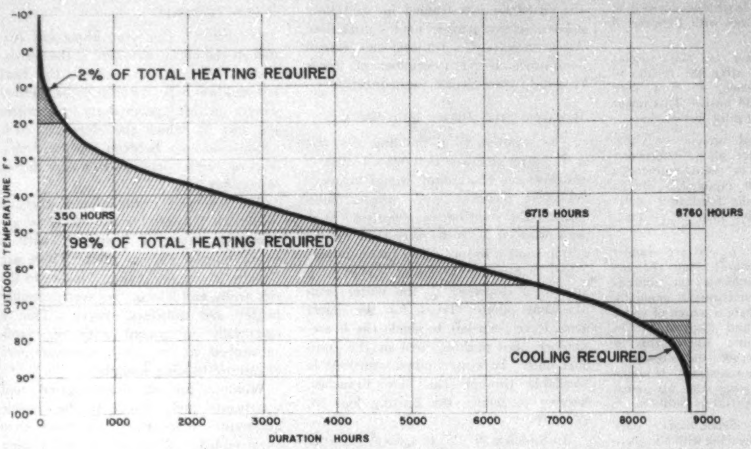


Fig. 2. Temperature duration curve

Heating-Cable Application and Selection

The heating cables buried in the ceilings have a capacity of about 60,000 Btu per hour at 240 volts and 53,000 Btu per hour at 220 volts. The manufacturer of the heating cables estimated the loss at nearer 48,000 Btu per hour and is confident that this will be proved through operation. From a practical point of view, if the installed capacity is excessive, nothing much is lost except a few installation dollars, and perhaps an occasional high, but explainable, demand. This is not serious in view of the pilot nature of the installation. On the other hand, if capacity were short, discomfort would certainly be experienced and the project could be considered a failure.

Table I. Temperature Duration, Rhode Island, July 1953 Through June 1954

Temperature, F	No. of Hours Below	Per Cent of Heating Load Below
100	8,760	
90	8,720	
80	8,450	
70	7,500	
65	6,715	100
60	5,825	81.5
50	4,175	44
40	2,475	20
35	1,660	12
30	1,025	7
25	600	4
20	350	2
15	197	1
10	71	
5	37	
0	14	
5	5	
10	0	

In this particular application, where direct-conversion equipment is involved in the same house with a heat pump, buried cables forming radiant ceilings were selected for several reasons:

1. Requires no wall space and otherwise is completely inconspicuous except for thermostats, thereby in no way restricting the arrangement of furniture or decor.
2. Represents a type of heating different in nature from the heat pump which in this case is forced air.
3. Features minimum cost which, at the same time, includes complete zoning and no expected maintenance.
4. In no way in conflict with the heat-pump installation.

Heat-Pump Application and Selection

For design purposes the forced-air system supplied by the heat pump is designed for 60,000-Btu-per-hour loss and 36,000-Btu-per-hour gain. In the selection of equipment, the water-source machine was selected because it appeared to be simpler than the air-source machines. For example, there is no "defrost cycle," and the available equipment is not so unitized in design as to almost forbid field maintenance. The water-source machine also offers better prospects in operating economies.

In the selection of size, manufacturer's specifications for a so-called 3-ton heat pump provide 36,000-Btu-per-hour heat absorption, and 42,000-Btu-per-hour on heating. A 5-ton machine provides 61,000-Btu-per-hour heat absorption and 70,000-Btu-per-hour on heating. Offhand, the 5-ton unit might seem desirable because it has more than enough capacity for both heating and cooling. However, this is not the case, because it is well known that oversizing on cooling is very undesirable. The machine will fulfill the cooling demand with short-cycled blasts of cold air, which, besides producing discomfort, may cause condensation of vapor to take place in the conditioned area rather than on the air coil of the heat pump. There is a resulting clamminess. Oversizing on heating results in similar short-cycling, which also impairs efficiency somewhat.

The conclusion reached was that, in order properly to size a heat-pump system for this climate, one of the following variations would probably be appropriate:

1. A dual heat pump, i.e., two units in tandem; one unit to operate for cooling with changeover to supply peak demand on heating, the other to operate for heating.
2. A single heat pump with direct-conversion heaters as boosters.
3. A single heat pump with storage capacity available for peak heating demand. Storage might be in water or salts.

Any of these might have been tried; however, one had to be chosen. The storage system was eliminated because of high first cost and as being too elaborate to be construed as "packaged."

Before proceeding further with the elimination, an average year was examined in order to determine the general nature of temperature duration. Referring to Table I, for a typical year there is tabulated the number of hours in the year when the temperature was below the values listed in the "Temperature" column. Since the temperature did not exceed 100 F, for example, there are 8,760 hours when the temperature was below 100 F. In further explanation, there were 6,715 hours when the temperature was below 65 F and 350 hours when it was below 20 F.

The third column in Table I lists the percentage of annual heating load occurring below the temperatures in the first column, e.g., only 2% of the total annual heating requirement occurs below 20 F. This means that if a basic heat pump covers heating losses to 20 F, then it supplies 98% of the total Btu's used, and the tandem or supplementary unit supplies 2% of the total use; see Fig. 2.

This indicates probable few hours of use for the second unit of a tandem heat pump, similarly for any supplementary direct-conversion heaters. Furthermore, there seemed to be more reason to believe that the heat loss for the house would be under 60,000 Btu per hour rather than over.

Direct-conversion heaters are available which contain, e.g., 12 1-kw units which can be connected in any combination to give a total output of from 1 to 12 kw. It seemed reasonable, therefore, to apply the flexibility of direct-conversion heaters to take care of the admittedly nebulous difference between heat-pump capacity and the actual house loss figure.

The resulting heat-pump installation utilizes a 3-ton machine in series with a duct heater containing 12 1-kw units. The duct heaters are tentatively connected in two banks of 3 kw each, each bank being controlled by an outdoor thermostat. In addition, one 3-kw bank is used for reheat in dehumidification.

Heat-Pump and Duct-Heater Installation and Operation

The installation and operation of the heating plant designed around the heat pump is best considered part by part, beginning with the water circuit.

WATER CIRCUIT

The electrical input to the well pump must be taken into account when computing the coefficient of performance of the heat-pump system. The pump itself should be correctly sized; the plumbing should develop as little resistance to flow as possible. If, at the same time, the pump can also be used for domestic purposes, this is much better.

The water requirement for the heat pump (50 F water) is about 9 gpm on heating and 1 to 2 gpm on cooling. To take care of this, a 1/2-horsepower jet pump is used, as shown in Fig. 3. In supplying the 9 gpm, the well pump delivers directly to the heat pump. However, this direct circuit cannot be used to deliver the 1 to 2 gpm for cooling because the pump is greatly oversized for this small flow, causing a continuous series of stops and starts. Water is therefore delivered from the storage tank for the cooling requirement.

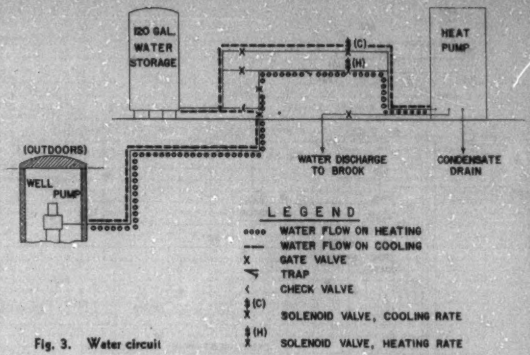


Fig. 3. Water circuit

A 120-gallon storage tank is supplied from the well pump through a check valve. This check valve prevents the storage tank from discharging into the heat pump (on heating), thereby maintaining satisfactory domestic pressure. The sizing of the storage tank is critical in the sense that its capacity must be adequate to domestic needs over any period of time that the heat pump runs. Whenever the heat pump shuts down, the well pump continues to run until storage tank pressure is recovered.

The results of delivering 9 gpm from the tank have been explored. Valves are arranged so that this can be done, although there is little practical value in this arrangement. Domestic pressure drops to below the useable range. A 1/2-horsepower pump will satisfy flow or pressure requirements in this instance, but not both.

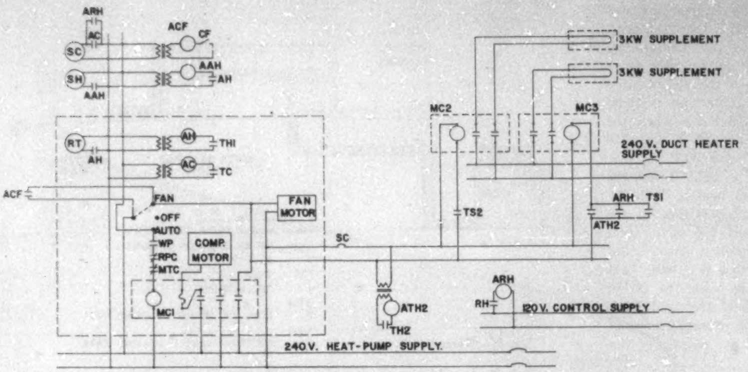
AIR DISTRIBUTION

Conditioned air is distributed through a duct system designed to handle the 1,200-cubic-feet-per-minute output of the heat pump. Ample outlets and returns are provided as shown in Figs. 1(B) and (C). In the bedroom areas, supplies are located high on the inside wall with returns low on the outside wall. This provides for flexible use of wall space and keeps furniture from hampering the supply air stream. Floor diffusers supply conditioned air to the family room-kitchen area, where they are located under the large glass-wall areas. One large return is located near the floor in the center of the area. The living-recreation room is in the basement where the duct trunks run near the main beam in the ceiling. In this area, ceiling diffusers are used for supplies, with returns low in the outside walls. Outdoor air can be introduced into the return air flow. An interesting sidelight is that the duct system would have had to handle over 2,000 cubic feet per minute if a tandem heat-pump plant had been selected, increasing the cost of duct work considerably.

HEAT PUMP AND SUPPLEMENTARY DUCT HEATERS

The electrical controls, shown schematically in Fig. 4, provide for automatic operation, including changeover from heating to cooling to dehumidification. The heat pump has a 4-way valve for automatic changeover, which is refrigerant actuated through the positioning of the throwover solenoid *RT*. A thermostat, with two heating stages *TH1* and *TH2*, one cooling stage *TC*, and a fan "auto-on" switch *CF* controls heating, cooling, and air circulation. When the first stage of heating *TH1* closes, the heating-water solenoid valve *SH* is energized and opens; at the same time, *RT* throws to heating. As soon as water is flowing through the heat pump, the compressor and circulating air fan start through the water pressure switch *WP*. The second stage of heating *TH2* provides quick pickup by switching in one 3-kw duct-heater combination

Fig. 4. Electrical schematic



LEGEND

MC1, MC2, MC3	MAGNETIC CONTACTORS	RH	RELATIVE HUMIDITY CONTROL (HUMIDISTAT)
MTC	COMPRESSOR MOTOR THERMAL CUTOUT	AH	FIRST AUXILIARY TRANSFORMER RELAY (HEATING)
RPC	REFRIGERANT PRESSURE CUTOUT	AAH	SECOND AUXILIARY TRANSFORMER RELAY (HEATING)
WP	WATER PRESSURE RELAY	AC	AUXILIARY TRANSFORMER RELAY (COOLING)
SH	SOLENOID (HEATING WATER)	ARH	AUXILIARY RELAY (RELATIVE HUMIDITY)
SC	SOLENOID (COOLING WATER)	ATH2	AUXILIARY TRANSF. RELAY (QUICK PICK-UP HEATING)
RT	REFRIGERANT THROWOVER SOLENOID	TS1	OUTDOOR THERMOSTAT, FIRST STEP SUPPLEMENTARY HEAT
TH1	INDOOR THERMOSTAT HEATING - STAGE 1	TS2	OUTDOOR THERMOSTAT, SECOND STEP SUPPLEMENTARY HEAT
TH2	INDOOR THERMOSTAT HEATING - STAGE 2	SC	SUPPLEMENTARY HEATER CUTOUT
TC	INDOOR THERMOSTAT COOLING		
CF	CIRCULATING FAN CONTROL FROM INDOOR THERMOSTAT		
ACF	AUXILIARY TRANSFORMER RELAY (CIRCULATING FAN)		

together with the heat pump. The cooling stage *TC* energizes the cooling-water solenoid *SC* valve, again starting the heat pump on water flow *WP*. Note that *RT* is normally in the cooling position. The auto-on switch *CF* is used to control the circulating air fan from upstairs so that it can be run continuously or when the compressor is running.

The heat pump as supplied is represented in Fig. 3 within the dashed lines. Included is a 3-position selector switch; "auto, "off," and "fan."

High humidity is controlled by a humidistat located in the living room; see Fig. 1(C). On a rise in relative humidity the humidistat energizes an auxiliary relay which starts the heat pump on cooling and switches in one 3-kw duct-heater combination. Return air from the room first passes over the cold coil in the heat pump, where its temperature is lowered and moisture condensed. The air then passes over the duct heater, which raises the temperature. This discharges the air to the living spaces at about the same temperature as the room return, but of course at a lower relative humidity.

It is possible for the humidistat and the thermostat to call for operation at the same time. When this happens on cooling, the operation of the system will first satisfy the setting on the humidistat before throwing over to cooling. It is not likely that heating and dehumidification will be required at the same time; however, if this should happen, then the heating requirement will first be satisfied before throwing over to dehumidification.

Two bulb-type thermostats with capillary tubes control the duct heaters on outdoor temperature, so that the electrical features can be mounted indoors and near the heat pump. The pickup setting is adjustable. To determine the temperature settings for these, the operational balance point of the heat pump is first determined.

Radiant-Ceiling Installation and Operation

The heating cables buried in plaster to form radiant ceilings are installed in accord with manufacturer's recommended procedures, except as mentioned later. In the case of two small lavettes, Figs. 1(B) and (C), the ceiling areas are too small to accommodate the smallest heating cable spaced at the 1.5 inch minimum. Baseboard heaters are used to heat these areas. Complete zoning is provided through thermostats located in each room and hallway. The storage and utility end of the basement is considered normally unheated. Space heaters are provided, however, to warm the area whenever necessary. Control of these heaters is manual.

Heating coils are supplied on reels and each wattage has a different length. A formula is used to convert the coil length and ceiling dimensions into a spacing distance. Apparently the spacing formula has a safety factor, the result being that the supply of wire was exhausted before the ceiling area in each case, causing no inconvenience.

In following the procedure for installing the cables, a staple gun was used to fasten the cables to the rock lath. Unfortunately the staple gun used was not the proper one: a wire guide had to be welded to the nose, spring tension was too great, and occasionally the gun would skip and fire with no staple. After the brown-coat plaster was applied and dried, a 600-volt Megger was used to test the insulation of each ceiling cable. In two rooms the insulation level was found to be below the 100,000 ohms specified as minimum by the manufacturer. In each case an ohmmeter was next used to determine where the insulation had been damaged. This was done by first connecting one leg of the ohmmeter to one of the leads of the heating cables at the thermostat box.

Fig. 5. Log sheets

RESIDENCE:
NAME ____________________
ADDRESS ____________________

DATE			TEMPERATURE				TIME				TOTAL HOUSE											REMARKS
MONTH	DAY	YEAR	MINIMUM	MAXIMUM	MEAN	° HOURS	METER READING	ELAPSED		SET CHART	BILLING METER KWH					SERVICE GRAPHIC					LOAD FACT.	
								HRS.	MIN.		BLACK X 10		RED X 10		TOTAL USE	KWH X 40			KW X 40			
											READ	USE	READ	USE		READ	DIF.	USE	READ	KW		

DATE			HEATING PLANTS																		
MONTH	DAY	YEAR	RADIANT CEILINGS						HEAT PUMP PLUS SUPPLEMENTAL I^2R												
			KWH X 20			KW X 20		L.F.	HT.P. KWH		I^2R KWH		WATER PUMP			USE ①+②+③	HEAT PUMP KW ①	I^2R KW ②	WATER PUMP KW ③	KW ①+②+③	L.F.
			READ	DIF.	USE	READ	KW.		READ	USE ①	READ	USE ②	CLOCK	HOURS USE	KWH ③						

Next, the other leg of the ohmmeter was connected to the blade of a broad pointed screw driver. In taking subsequent ohmmeter readings, the screw driver was used to make firm contact with the brown coat by enough pressure to penetrate the brown coat approximately 1/32 inch. Readings then were recorded as the ceiling was crisscrossed in order to determine the point of least resistance, and thereby locate the fault. It was found that the point of least resistance was approximately 80% of the highest value; however, the resistance gradient adjacent to the trouble spot was great. In each case a staple driven too hard against the insulation was the cause of insulation failure, and following the foregoing procedure each was located within a radius of an inch from the point of indicated least resistance. Plastic insulating tape was used to renew the insulation. After the ceilings were finished, all ceiling cables were tested again with the Megger and found to be all right.

Electrical Service and Instrumentation

ELECTRICAL SERVICE

An aerial 200-ampere 120/240-volt service supplies the electrical needs of the house through facilities in the service room, shown as a detail of Fig. 1(C). Distribution and subdistribution are through quick-lag breaker panels which, beside being simple and adequate, avoid the cluttered and commercial appearance of a series of fused safety switches, or the like, all supplied through troughs. Except that special metering is involved in this installation, a service room would probably be unnecessary.

INSTRUMENTATION

The electrical service is metered through a 2-dial off-peak meter for billing purposes. In addition, the total service is chart-metered on 15-minute kilowatt demand. The ceiling subpanel, the heat pump, and the duct heaters are metered separately for kilowatt-hour and 15-minute kilowatt-hour demand (maximum indicating). The water-pump use is estimated from the hours operation of the heat pump.

Separate metering of other appliances is believed to be unnecessary, since much experience data are available on the nature of the residential load without heating. Interest is confined to the nature of each type of electric heating and its affect on the total residential demand and use. The total data indicate the impact that this type of house can be expected to have on the utility's distribution system.

In addition, various nonelectrical quantities can also be recorded:

Water Circuit. Pressure gages are located on the water inlet and discharge connections to the heat pump. Setting of the pressure switch *WP* can be checked and, in addition, water flow through the machine can be estimated from the indicated drop of pressure through the water coil.

Immersible-bulb dial-type thermometers are also located at the inlet and outlet to show the temperature change in the water coil. A stop watch and 10-gallon milk can are used to meter the flow of water at discharge.

Air Circuit. Indicating column-type thermometers are used to measure the temperature at three points: 1. return room air to the heat pump, 2. air from the heat-pump air coil, and 3. beyond the duct heaters. Temperature pickup across the heat-pump air coil and the duct heaters can thereby be recorded and compared to show an indicated coefficient of performance.

Indoor and Outdoor Air. Temperature and humidity indoors and outdoors are recorded on chart-type instruments.

Heat Pump. Refrigeration pressure gages are used to show the head and suction pressures of the refrigerant gas at the compressor. Readings from these can be combined to form an indication of heat-pump performance.

This instrumentation is adequate to show the performance of the various electrical heating features, to check the output specification of the heat pump, and to check the heat-gain and heat-loss calculations.

RECORDS

A daily diary is kept up to date on equipment operation and performance. For convenience, some of this information is kept in log-sheet form as shown in Fig. 5.

At the writing of this paper, both heating plants have been operated satisfactorily. However, a performance report will not be released until substantial data are accumulated and evaluated.

Voltage Modulation on Aircraft Power Systems

R. E. KLOKOW
MEMBER AIEE

C. F. YOHE
ASSOCIATE MEMBER AIEE

Synopsis: This paper derives methods for calculating voltage modulation for a given speed disturbance with closed-loop voltage control. The magnitude of voltage modulation is quantitatively related to the open-loop frequency response of the voltage-regulating system.

VOLTAGE MODULATION may be defined in a broad sense as any change in the envelope of the a-c generator terminal voltage. For the case of sinusoidal voltage modulation, a precise definition is given in Appendix I.

Present and proposed military specifications are given in Appendix I. Present specifications limit voltage modulation of aircraft electric power systems to 2%. Proposed specification limits are in a range of 1.0% to 0.10%. The specifications do not explicitly state the character of voltage modulation (sinusoidal, periodic, or random), the frequency range, or the speed modulation of the generator rotor. It is hoped that the results of this paper will give information which can be used to clarify the specifications.

Since magnitudes in the order of 1.0% are involved, it is imperative that methods of calculating voltage modulation be available. It is equally important that insight into the problem of voltage modulation be gained. The causes and factors must be thoroughly investigated. Test results on actual systems are necessary, but are not sufficient to these ends. This paper presents results of part of a study of the voltage modulation problem. Before stating the precise problem considered in this paper, the broad causes of voltage modulation will be reviewed.

The generated terminal voltage is proportional to the product of flux and speed. Variations in either flux or speed will cause variations in generated voltage. Flux variations, and consequently voltage variations, may be produced by a-c generator excitation voltage changes, variation in brush contact drop (if brushes are used), nonsymmetrical generator construction, and load disturbances. It has been found from testing that the effect of variation in brush contact drop and nonsymmetrical generator construction on terminal voltage is negligible on well-designed generators. The effect of load disturbances are neglected in this paper and the generator is assumed to be operating with a balanced static load. Speed variations may be produced by torque pulsations of the prime mover, the genera-

---

Paper 57-470, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7-9, 1957. Manuscript submitted January 30, 1957; made available for printing May 17, 1957.

R. E. KLOKOW and C. F. YOHE are with the Westinghouse Electric Corporation, Lima, Ohio.

The authors wish to acknowledge the work of Prof. J. L. Klingenberger, Ohio Northern University, in measuring the open-loop transfer function of the example system, and of R. I. Daniels, mechanical design engineer, Westinghouse Electric Corporation, Lima, Ohio, for the construction of the special test equipment.

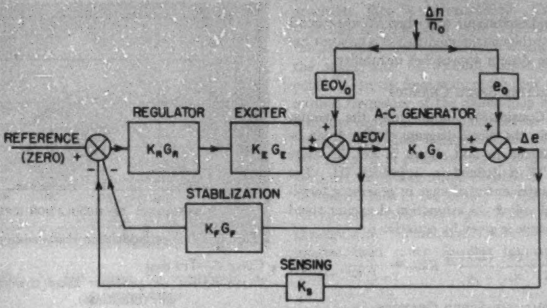


Fig. 1. Block diagram of typical voltage regulator system

tor drive, gearing, or load-switching transients. These pulsations may be amplified by the mechanical system depending on the frequency characteristics of any particular configuration. The characteristics of the mechanical system are vitally important to the problem of voltage modulation, but are not directly considered in this paper. The starting point of the study made here is with given speed modulation of the generator rotor.

Nomenclature

e = envelope of a-c generator-terminal voltage
ϕ = instantaneous flux per pole
n = instantaneous generator rotor speed (rpm)
ΔN = amplitude of sinusoidal speed modulation
f = frequency of speed changes
$\omega = 2\pi f$ = frequency of speed change in radians per second
KG = open-loop transfer function of complete system where G is a function of $(j\omega)$
$\Delta e'$ = change in e due to a-c generator speed changes
$\Delta e''$ = change in e due to exciter speed changes
Δe = instantaneous change in envelope of a-c generator terminal voltage
ΔE = maximum value of the sinusoidal change in the envelope of the a-c generator terminal voltage
K_RG_R = transfer function of regulator, where G_R is a function of $(j\omega)$
K_EG_E = transfer function of exciter, where G_E is a function of $(j\omega)$
K_GG_G = transfer function of a-c generator, where G_G is a function of $(j\omega)$
K_FG_F = transfer function of feedback circuit, where G_F is a function of $(j\omega)$
K_s = sensing circuit gain
$(EOV)_0$ = exciter output voltage at operating point
α = phase angle of flux variations with respect to speed variations

Subscript 0 indicates value of variable at operating point

Statement of Problem

The following discussion is limited to a-c generators with sinusoidal disturbances and a linearized system. However, the equations and definitions developed could be adapted for d-c generators, any given speed disturbance, and nonlinear systems. The envelope of the a-c terminal voltage is given by

$$e = K\phi n \tag{1}$$

where

ϕ = space fundamental flux per pole
n = instantaneous generator rotor speed, rpm
K = a constant of proportionality

Voltage modulation may be produced by either a flux change, a speed change, or a combined change of both variables. Equation 2, developed in Appendix II for the case of constant flux and sinusoidal variations in speed, is

$$\text{Per-cent voltage modulation} = \frac{\Delta N}{n_0} \times 100 \tag{2}$$

Therefore, with sinusoidal speed disturbances and constant a-c generator field flux, the per-cent voltage modulation is equal to the percentage speed variation.

Equation 3 developed in Appendix II for the case of constant speed and sinusoidal flux variations is

$$\text{Per-cent voltage modulation} = \frac{\Delta\Phi}{\phi_0} \times 100 \tag{3}$$

This equation states that for constant speed and sinusoidal variations in flux the percentage voltage modulation is equal to the percentage variation of flux. Equation 4 developed in Appendix II for the case of variable flux and variable speed, is as given in the following equation 4:

$$\begin{matrix}\text{Per-cent voltage}\\ \text{modulation}\end{matrix} = \left|\frac{\Delta N}{n_0}\underline{/0} + \frac{\Delta\Phi}{\phi_0}\underline{/\alpha}\right| \times 100 \tag{4}$$

If both speed and flux are sinusoidally varied, the per-cent voltage modulation is the magnitude of the phasor sum of the change due to each of the variables.

When an automatic voltage regulator is controlling the a-c generator terminal voltage, and speed disturbances are applied to the rotor, an error in terminal voltage will be seen by the regulator. The regulator will attempt to make corrections for the error in terminal voltage. These changes produced by the regulator will eventually be seen as changes in a-c generator field flux. Thus there is a combined variation in both speed and flux and the resulting voltage modulation will be given by equation 4. However, magnitude and phase angle of the flux change with reference to the speed change is not known since it is dependent upon the frequency response characteristics of the control system. It is the objective of this paper to derive methods of determining the resulting voltage modulation under this condition.

Assumptions

The following assumptions are made in order to expedite the discussion of this subject. However, some of these assumptions will be removed in the development of the paper:

1. The system has been linearized about no-load operation.
2. Only small excursions of the variables are considered.
3. Speed disturbances are sinusoidal functions.
4. Voltage variations due to changes in brush contact drop, load disturbances, and nonsymmetrical machine construction are neglected.
5. Only single-generator operation is considered.
6. The sinusoidal speed changes are assumed to be that of the generator rotor and exciter, and the influence of the mechanical system is not considered.
7. Effect of magnetic amplifier supply voltage variation on voltage modulation is neglected.

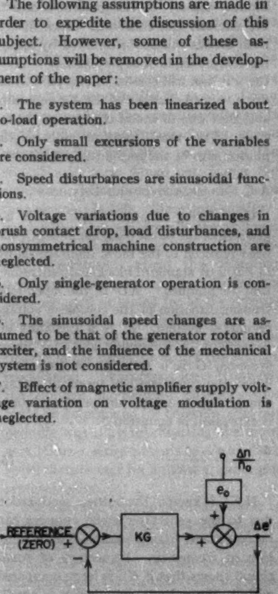


Fig. 2. Simplified block diagram of voltage regulator system with speed inputs to generator only

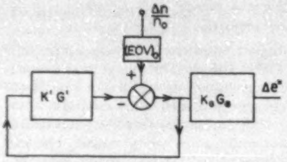


Fig. 3. Simplified block diagram of voltage regulator system with speed inputs to exciter only

Solution

Although the method presented is general, a particular system will be used in the description. Fig. 1 shows a block diagram of a typical voltage regulator system. The system is specific in the arrangement of the components but general in the sense that each component could have any transfer function. With this block diagram given, the problem reduces to that of determining

$$\frac{\Delta e/e_0}{\Delta n/n_0}$$

Since in the system there are two speed inputs, namely, to the generator and to the exciter, the combined effect due to both inputs will be determined by applying the principle of superposition. Thus, the solution will consist of two parts, that due to speed changes of the a-c generator and that due to speed changes of the exciter. The total solution will be the phasor sum of the individual solutions.

A-C Generator Speed Changes

Considering only the input to the a-c generator, the block diagram of Fig. 1 has been reduced to the block diagram shown in Fig. 2 by standard block diagram reduction techniques.[1] KG is the open-loop transfer function of the complete voltage regulator system. K is a constant independent of frequency and, for linear systems, G is a function of frequency. Thus, the component of terminal voltage change due to generator speed variations is expressed in equation 5

$$\frac{\Delta e'}{e_0}=\frac{\Delta n}{n_0}\,\frac{1}{1+KG} \tag{5}$$

If KG is known, the voltage modulation due to the generator speed changes can be computed from equation 5 as a function of frequency. Inspection of equation 5 shows that if KG is large compared with 1, the voltage modulation is small; if KG approaches zero the voltage modulation with closed-loop control will be equal to that of a generator with constant field excitation; and when KG approaches -1 the voltage modulation is greatest but the system approaches instability.

Exciter Speed Changes

Considering inputs into the exciter only the block diagram of Fig. 1 can be reduced to that shown in Fig. 3, where $K'G'$ is defined in Appendix III. The component of change in generator terminal voltage as a function of exciter speed change is given by equation 6:

$$\frac{\Delta e''}{e_0}=\frac{\Delta n}{n_0}\,\frac{(EOV)_0}{e_0}\,K_GG_G\,\frac{1}{1+K'G'} \tag{6}$$

Combined Speed Changes

The total change in terminal voltage due to the effect of exciter and generator speed changes is

$$\frac{\Delta e}{e_0}=\frac{\Delta e'}{e_0}+\frac{\Delta e''}{e_0}=\frac{\Delta n}{n_0}\left[\frac{1}{1+KG}+\frac{(EOV)_0}{e_0}\,K_GG_G\,\frac{1}{1+K'G'}\right] \tag{7}$$

Expressed in terms of a ratio of per-cent voltage modulation to per-cent speed modulation, equation 7 becomes

$$\frac{\%\text{ voltage modulation}}{\%\text{ speed modulation}}=\frac{1}{1+KG}+\frac{(EOV)_0}{e_0}\,K_GG_G\,\frac{1}{1+K'G'} \tag{8}$$

For the systems investigated by the authors the second term of equation 8, which is the voltage modulation due to the exciter speed changes has been small in comparison to the voltage modulation due to the generator speed changes. Neglecting this term, equation 8 simplifies to equation 9:

$$\frac{\%\text{ voltage modulation}}{\%\text{ speed modulation}}=\frac{1}{1+KG} \tag{9}$$

Equation 9 shows the simplicity of the method of determining percentage voltage modulation in terms of the given percentage speed modulation. The only information required is the open-loop transfer function of the complete voltage regulator system which can be calculated or measured. This information is available in the design since it is used for determining the stability of voltage regulator systems in linear regions.

Verification of Analytical Method

Curve A of Fig. 4 represents the results of a test on a 40-kva generator system regulated by a static-type magnetic-amplifier voltage regulator. These test data were obtained by a special test fixture, as shown in Fig. 5. The a-c generator rotor was driven at a constant speed by a dynamometer. The relative speed was modulated sinusoidally by vibrating the a-c stator which was mounted in trunnion bearings. The amplitude of speed modulation was measured by providing constant excitation to the a-c generator, and measuring per-cent voltage modulation which is equal to the percentage speed modulation. With the same speed modulation, voltage modulation was again measured with the a-c generator controlled by the regulator. The ratio of voltage modulation to speed modulation was then determined by dividing the second measurement by the first for various modulation frequencies.

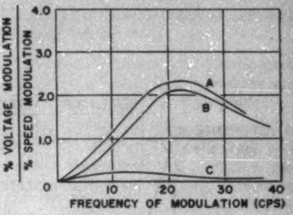


Fig. 4. Voltage modulation characteristics

Curve A—Test data
Curve B—For a-c generator speed changes only (calculated)
Curve C—For exciter speed changes only

Fig. 6 shows the open-loop frequency response characteristics of the 40-kva voltage regulator system. This data was obtained by frequency response measurements and calculations, and the information used in equation 9 to determine the voltage modulation characteristics and plotted as curve B of Fig. 4.

The computed results compare favorably with the test curve. The difference between the two curves (A and B) is due to the inaccuracy in test measurements of voltage modulation in the range of 0.1% to 1.5%, the neglect of speed input to the exciter and errors in obtaining the open-loop frequency response.

The magnitude of voltage modulation due to the speed variation of the exciter is shown as curve C of Fig. 4. This effect is seen to be small in terms of magnitudes, but has even less influence when added phasorially to curve B.

Interpretation of Results

Observation of Figs. 4 and 6 and inspection of equation 9 lead to the following interpretation:

When $|KG|$ is much greater than unity, equation 9 reduces to the following.

Fig. 5. Test fixture for modulating a-c generator speed

$$\frac{\%\ \text{voltage modulation}}{\%\ \text{speed modulation}} \approx \frac{1}{KG}$$

When $|KG|$ is much less than unity, equation 9 reduces to

$$\frac{\%\ \text{voltage modulation}}{\%\ \text{speed modulation}} \approx 1$$

When the magnitude of KG approaches -1, neither approximation holds. Fig. 6 shows that for a typical voltage regulator system the magnitude of KG is greater than unity at the low frequencies and less than unity at the high frequencies. Thus, it can be concluded that the ratio of per-cent voltage modulation to per-cent speed modulation will be less than unity at the low frequencies and equal to unity at the high frequencies. The ratio will be greater than unity at and near the natural frequency of the closed-loop voltage regulator system. The amplification in the region of the closed-loop natural frequency could reach extremely high values if the angle of KG should approach, 180 degrees when the magnitude of KG approached unity. However, for a 45-degree phase margin design, the angle of KG is held below 135 degrees when the magnitude of KG is equal to one. Consequently, for a well-designed voltage regulator system, the peak amplification can be held below an amplification of three depending upon other specification requirements, such as steady-state regulation and recovery time.

Physically, the three regions of the voltage modulation curves of Fig. 4 can be interpreted as follows: At low frequencies the regulator can change the flux in the a-c generator in the proper manner to correct for the changes in terminal voltage caused by the disturbing speed inputs. At the very high frequencies the correcting signals sensed by the voltage regulator are attenuated in magnitude to such a degree that very little change in the a-c generator flux is accomplished. Consequently at these very high frequencies the per-cent voltage modulation would be the same with or without closed-loop control. In the frequency range near the natural frequency of the closed-loop system, the regulator tries to correct for the changes in terminal voltage but the correction signal is in phase or nearly in phase with the disturbing signal because of the 180-degree shift in the sensing circuit plus the approximately 180-degree phase shift through KG. Consequently, the ratio is greater than unity in this frequency range.

Correlation of Voltage Modulation Characteristics with Other System Requirements

The voltage modulation characteristic is definitely related to the steady-state and transient performance of the voltage regulator system when subjected to load switching and speed disturbances. It is not the purpose of this paper to express these relationships quantitatively, but to give a qualitative discussion to show that the relationships exist.

Per-cent voltage regulation is equal to

$$\frac{1}{1+K}\times 100$$

K, the steady-state open-loop gain, affects voltage modulation, as can be seen from equation 9. Consequently, there is a definite relationship between voltage modulation and steady-state regulation.

Nyquist's stability criterion states that the locus of the vector $KG\ (j\omega)$ must not pass through the -1 point or encircle this point for a closed-loop system to be stable. Several textbooks show the relationship between the open-loop response $KG\ (j\omega)$ and the stability of servomechanism and regulator systems. Other authors have shown the relationship between frequency response and transient response. For purposes of this paper, it will suffice to say that there is a relationship between the stability of the system and the voltage modulation characteristic. In general, it may be said that the more oscillatory a closed-loop system is to load changes, the higher the ratio of per-cent voltage modulation to per-cent speed modulation will be at the natural frequency of the closed-loop system.

There is also a relationship between steady-state voltage regulation and stability, consequently, all three system requirements are interrelated, i.e., improvement in one (such as per-cent regulation) may result in making the other two (e.g., voltage modulation and stability) worse.

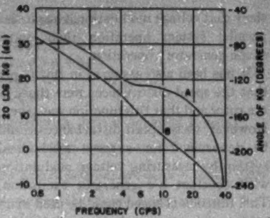


Fig. 6. Open-loop frequency response characteristics of voltage regulator system

Generalization of Method

Several assumptions were enumerated at the start of the paper. Most of these were made in order to confine the discussion and the particular system treated to a defined region. The discussion has been limited to no-load operation which is usually the most severe condition for voltage modulation. However, other load conditions could be treated by determining a new KG for the desired load.

Only small excursions of the variables were considered in order to linearize for saturation of the control system components and avoid the nonlinearity of the second-order term as given in equation 20. This second-order term will be negligible for all practical systems where the voltage modulation is small. For a given case where the speed modulation is 1.0% and the percentage change of a-c generator flux is 1.0%, to include the second-order term would not change the computed voltage modulation by more than one part in 100. Since only small excursions are encountered in the voltages of the various components for this problem, linearization about the operating point gives accurate results. Constants are computed or tested for at the operating point under consideration.

Only sinusoidal speed disturbances were treated in the development of the voltage modulation characteristics. However, the voltage modulation characteristics could be determined for nonsinusoidal but periodic waves by representing the input as a Fourier series or for a given random speed disturbance by an analog computer representation.

The effect of voltage modulation due to brush drop and nonsymmetrical machine construction was neglected. However, it has been found on all systems investigated that these effects are negligible on well-designed machines.

Only single-generator operation has been considered; however, test results show that voltage modulation is less under parallel system operation than under single-generator operation.

It has been also assumed in this paper that the speed disturbances were those of the rotor and that this function was given. However, these speed disturbances could be determined for a given mechanical system by measuring voltage modulation with constant a-c generator excitation. This function of speed could be used in an analog computer simulation to determine voltage modulation for a particular closed-loop system.

The effect of the variation of the supply voltage to the magnetic amplifiers on voltage modulation could be evaluated by including extra loops in the block diagram. However, the calculations become lengthy and it has been found that these effects are negligible.

Conclusions

The ratio of per-cent voltage modulation to per-cent speed modulation has been quantitatively related to the open-loop frequency response characteristic of the voltage regulator system. The solution is stated mathematically in equation 9, and is discussed in a previous section. Applying sinusoidal speed disturbances to the generator of a closed-loop voltage regulator system results in the following ratio of per-cent voltage modulation to per-cent speed modulation: 1. At frequencies under the natural frequency of the closed-loop system, this ratio is less than one; 2. over natural frequency it is equal to unity; and 3. at and near the natural frequency, it is greater than one.

From the results of this paper it is apparent that per-cent voltage modulation cannot be specified independently of per-cent speed modulation, i.e., before voltage modulation can be calculated, the speed modulation must be known. On a particular system, the voltage modulation could be objectionable when the generator is driven by a given test stand and be within specification limits when driven on an aircraft. Although this paper solves the problem of determining voltage modulation with given generator rotor speed, determining voltage modulation on an actual system necessitates a study of the complete system. Cognizance must be taken of the drive, the generator mechanical system, the gearing between drive and generator, the frequency controller of the drive, and the speed or torque pulsations into the drive from the prime power source.

Since the ratio of per-cent voltage modulation to per-cent speed modulation is interrelated with the steady-state regulation of the system and the transient performance of the system, no one requirement can be looked at individually but rather they must all be considered simultaneously.

Appendix I. Definition of Voltage Modulation and Military Specifications

The following is a summary of the voltage modulation requirements from some of the military specifications. There is a trend in the later specifications to decrease the voltage modulation limits. To comply with these more stringent specification requirements, more extensive consideration must be given to the magnitude and frequency of speed variations and to the design of the closed-loop frequency characteristics of the control system. However, these specifications do not explicitly state speed variation characteristics.

Military Specifications

1. *MIL-G-6099 (April 1950)*

4.5.11. Output Voltage Modulation. The output voltage modulation shall be determined by means of a calibrated oscilloscope or other satisfactory means when the generator is operated at minimum, average, and maximum rated speed both at no load and full load. The per-cent modulation, computed as indicated below, with the regulator in proper operation, shall not exceed 1 per cent plus that obtained when the generator is controlled by a hand rheostat, and in no case shall exceed 2 per cent. Per-cent modulation shall be computed as follows:

$$\text{Per-cent modulation}=\frac{E_{\max}-E_{\min}}{E_{\max}+E_{\min}}\times 100$$

2. *MIL-G-6099A(ASG) (November 27, 1956)*

4.5.13. Output Voltage Modulation. The output voltage modulation shall be determined when the generator is operated at minimum, average, and maximum rated speeds both at no load and rated current at rated voltage. The per-cent modulation, computed as indicated below, with the regulator in proper operation shall not exceed 1 per cent. Per-cent modulation shall be computed as follows:

$$\text{Per-cent modulation}=\frac{E_{\max}-E_{\min}}{E_{\max}+E_{\min}}\times 100$$

3. *MIL-G-7894A (ASG) (May 17, 1955)*

3.2.1.2.6. Amplitude Modulation. The amplitude modulation of the voltage wave shall not exceed 2 per cent at any frequency.

4. *XEL-907 (June 1, 1956)*

3.1.3.2.5. Voltage Modulation.

3.1.3.2.5.1. Amplitude. The modulation of voltage shall not exceed an amplitude of 3.3 volts (115 volt, line-to-neutral RMS system) as measured peak-to-peak difference between the minimum voltage reached and the maximum voltage reached during any one second.

3.1.3.2.5.2. Frequency Characteristic of the Voltage Modulation. The frequency characteristics of the voltage modulation shall be confined within the envelope of limits shown in Fig. 2. [Fig. 7 of this paper]

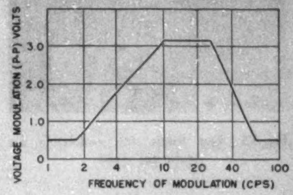


Fig. 7. Frequency characteristics for voltage modulation in a-c systems (Fig. 2, ref. 7)

Definition of Voltage Modulation

Voltage modulation has been defined in the listed military specifications in terms of percentage and peak-to-peaks volts. Per-cent voltage modulation is used throughout this paper in order to make the definition of voltage modulation independent of line-to-line, line-to-neutral, peak-to-peak, and peak volts. If consistent units of voltage are used in the following, the percentage voltage modulation computed will be correct:

$$\begin{array}{c}\text{Per-cent voltage}\\ \text{modulation}\end{array}=\frac{E_{\max}-E_{\min}}{E_{\max}+E_{\min}}\times 100 \qquad (10)$$

Simplified Equation for Computing Per-Cent Voltage Modulation

The voltage $E_{\max}$ and $E_{\min}$ are graphically presented in Fig. 8 together with e_0 and ΔE for sinusoidal voltage modulation. An equivalent expression for per-cent voltage modulation in terms of e_0 and ΔE is

$$\begin{array}{c}\text{Per-cent voltage}\\ \text{modulation}\end{array}=\frac{E_{\max}-E_{\min}}{E_{\max}+E_{\min}}\times 100$$

$$=\frac{(e_0+\Delta E)-(e_0-\Delta E)}{(e_0+\Delta E)+(e_0-\Delta E)}\times 100$$

$$\begin{array}{c}\text{Per-cent voltage}\\ \text{modulation}\end{array}=\frac{\Delta E}{e_0}\times 100 \qquad (11)$$

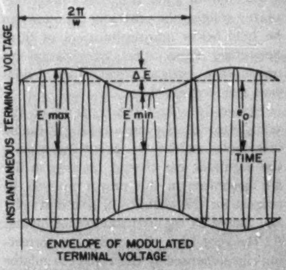

2π/w
INSTANTANEOUS TERMINAL VOLTAGE
ΔE
E max
E min
e_0
TIME
ENVELOPE OF MODULATED TERMINAL VOLTAGE

Fig. 8. Modulated a-c generator output voltage

Random-Type Voltage Modulation

The discussion in this paper has been limited to voltage modulation of a sinusoidal nature. If speed disturbances are of a random nature, the voltage modulation would be of a random nature. The present definition of voltage modulation has little meaning for the random type of speed disturbances. The voltage modulation measured from the minimum dips to the maximum peak will depend upon at what interval of time the measurements are taken and the frequency response characteristics of the metering devices. The present definition of voltage modulation will not necessarily give numerical values which are indicative of the effect of random voltage disturbances on utilization equipment.

Appendix II. Derivation of Equations

The purpose of this appendix is to derive expressions for the a-c generator terminal voltage for three combinations of generator air-gap flux and generator rotor speed. The three cases considered are

1. Constant flux and sinusoidal speed variations.
2. Constant speed and sinusoidal flux variations.
3. Sinusoidal variation of both flux and speed.

An expression for the envelope of the a-c generator terminal voltage is

$$e=K\phi n \tag{1}$$

Constant Flux and Sinusoidal Speed Variations

An expression for the change in the envelope of the a-c generator terminal voltage will be derived for the case of constant flux and sinusoidal speed variations. Let

$\phi=\phi_0$, a constant

$n=n_0+\Delta N\sin\omega t$

$e=e_0+\Delta e$

Substituting in equation 1 gives the desired result

$$e_0+\Delta e=K\phi_0\Delta N\sin\omega t+K\phi_0 n_0 \tag{12}$$

Since $e_0=K\phi_0 n_0$

$$\Delta e=K\phi_0\Delta N\sin\omega t \tag{13}$$

Dividing both sides of this equation by e^0 and assuming that e_0 is rated voltage gives

$$\frac{\Delta e}{e_0}=\frac{\Delta N}{n_0}\sin\omega t$$

Expressing the results in terms of the magnitude of the sinusoidal envelope gives

$$\frac{\Delta E}{e_0}=\frac{\Delta N}{n_0}\text{ or}$$

$$\text{per-cent voltage modulation}=\frac{\Delta N}{n_0}\times 100 \tag{2}$$

or, stated in words, per-cent voltage modulation equals per-cent speed modulation for the case of constant flux and sinusoidal speed perturbations.

Constant Speed and Sinusoidal Flux Changes

Let

$n=n_0$, a constant

$\phi=\phi_0+\Delta\Phi\sin\omega t$

$e=e_0+\Delta e$

Substitution of these relations in equation 1 yields

$$e_0+\Delta e=Kn_0\phi_0+Kn_0\Delta\Phi\sin\omega t \tag{14}$$

Using a similar approach to that used in the preceding section gives

$$\text{Per-cent voltage modulation}=\frac{\Delta\Phi}{\phi_0}\times 100 \tag{3}$$

Or per-cent voltage modulation equals per-cent flux modulation for the case of constant generator speed and sinusoidal flux variations.

Sinusoidal Variation of Both Flux and Speed

Let

$$n=n_0+\Delta n\sin\omega t \tag{15}$$

$$\phi=\phi_0+\Delta\Phi\sin(\omega t+\alpha) \tag{16}$$

Substitution into equation 1 gives

$$\begin{aligned}e_0+\Delta e&=K[\phi_0+\Delta\Phi\sin(\omega t+\alpha)]\times(n_0+\Delta N\sin\omega t)\\&=K\phi_0 n_0+K\phi_0\Delta N\sin\omega t+Kn_0\Delta\Phi\times\sin(\omega t+\alpha)+K\Delta\Phi\Delta N\sin(\omega t+\alpha)\sin\omega t\end{aligned} \tag{17}$$

Neglecting the second-order term and making use of the fact that $e_0=K\phi_0 n_0$

$$\Delta e=K\phi_0\Delta N\sin\omega t+Kn_0\Delta\Phi\sin(\omega t+\alpha) \tag{18}$$

Dividing both sides by e_0:

$$\frac{\Delta e}{e_0}=\frac{\Delta N}{n_0}\sin\omega t+\frac{\Delta\Phi}{\phi_0}\sin(\omega t+\alpha) \tag{19}$$

Writing equation 18 in terms of phasors gives

$$\text{Per-cent voltage modulation}=\left|\frac{\Delta N}{n_0}\underline{|0}+\frac{\Delta\Phi}{\phi_0}\underline{|\alpha}\right|\times 100 \tag{4}$$

Equation 4 states that the amplitude of the voltage modulation envelope is equal to the magnitude of the phasor sum of per-unit speed and per-unit flux for the case of sinusoidal speed flux changes.

Appendix III. Block Diagram Representation

Linearization of Product of Flux and Speed

The block diagram given in Fig. 1 shows that terminal voltage is a function of the sum of flux and speed changes. Derivation of the equation showing this relationship follows:

From equation 1

$$\left.\begin{aligned}&e=K\phi n\\&\text{Let: }\phi=\phi_0+\Delta\phi\\&n=n_0+\Delta n\\&e=e_0+\Delta e\end{aligned}\right\} \tag{1}$$

Substituting these in equation 1, yields

$$e_0+\Delta e=K(\phi_0+\Delta\phi)(n_0+\Delta n)$$

$$e_0+\Delta e=K(\phi_0 n_0+n_0\Delta\phi+\phi_0\Delta n+\Delta\phi\Delta n) \tag{20}$$

Neglecting the second-order term gives

$$\Delta e=Kn_0\Delta\phi+K\phi_0\Delta n \tag{21}$$

Dividing by e_0 gives

$$\frac{\Delta e}{e_0}=\frac{\Delta\phi}{\phi_0}+\frac{\Delta n}{n_0} \tag{22}$$

Representation of equation 22 is shown in the block diagram of Fig. 1, as is also speed input to the exciter, which was handled in a similar manner.

Incremental Variables

All variables used in the block diagrams are incremental variables. The constants are determined at the operating points of each component. For example, the gain and time constant of the generator are proportional to the slope of the generator saturation curve at the operating point. Use of incremental variables is justified in the study of voltage modulation because the changes are in the order of 1.0%. However, if the transfer characteristics are not linear over the range of variables considered, this nonlinear problem can be solved on an analog computer.

Reduction of Block Diagram

Rather than present step-by-step reduction of the block diagram shown in Fig. 1 to those given in Figs. 2 and 3, KG and $K'G'$ are defined by the following equations:

$$KG=\left(\frac{K_RG_RK_EG_E}{1+K_FG_FK_RG_RK_EG_E}\right)K_SK_GG_G \tag{23}$$

$$K'G'=K_RG_RK_EG_E(K_FG_F+K_SK_GG_G) \tag{24}$$

References

1. PRINCIPLES OF SERVOMECHANISMS (book), G. S. Brown, D. P. Campbell. John Wiley & Sons, Inc., New York, N.Y., 1948.

2. SERVOMECHANISMS AND REGULATING SYSTEM DESIGN, VOL. I (book), H. Chestnut, R. W. M. Mayer. John Wiley & Sons, Inc., 1955.

3. DEVELOPMENT OF DEFINITIONS FOR VOLTAGE MODULATION CHARACTERISTICS IN 380- TO 420-CYCLE-PER-SECOND AIRCRAFT ELECTRIC SYSTEMS, O. Markowitz. *AIEE Transactions*, vol. 75, pt. II, Sept. 1956, pp. 193–97.

4. GENERAL SPECIFICATION FOR AIRCRAFT ALTERNATING CURRENT GENERATORS AND REGULATORS. *Military Specification MIL-G-6099*, Apr. 19, 1950.

5. GENERAL SPECIFICATION FOR AIRCRAFT ALTERNATING CURRENT GENERATORS AND REGULATORS. *Military Specification MIL-G-6099A(ASG)*, Nov. 27, 1956.

6. CHARACTERISTICS OF AIRCRAFT ELECTRIC POWER. *Military Specification MIL-G-7894A (ASG)*, May 17, 1955.

7. GENERAL SPECIFICATION FOR ALTERNATING CURRENT, CONSTANT FREQUENCY, SINGLE GENERATOR, GENERATING SYSTEM. *Military Specification XEL 907*, June 1, 1956.

Design of Generating Equipment for Electra Turboprop Aircraft

J. R. M. ALGER
ASSOCIATE MEMBER AIEE

W. E. WARDEN
ASSOCIATE MEMBER AIEE

W. O. HANSEN
NONMEMBER AIEE

L. KLEIN
ASSOCIATE MEMBER AIEE

WITH the advent of the Electra turboprop aircraft, a new phase was introduced in the type of electric power generated and distributed on commercial aircraft. Previously to this time, nearly all commercial aircraft used a 28-volt d-c system. However, the requirement for higher rated systems and the requirement of the load equipment for a-c power resulted in the Electra using the 115/200-volt 3-phase Y-connected a-c system. Many models of military aircraft had already used a-c systems but the systems in production at that time mainly used rotating exciter generators with sensitive protection relays. At the time that the Electra design was initiated a static excitation system with magnetic amplifier controls and protection was under development. This system was developed in order to increase reliability and provide closer limits of system performance. This static excitation system has been applied to the Electra and this paper discusses its design features.

The system described here uses basic designs described in a paper by Plette and Butler.[1] This paper is confined to describing the designs peculiar to the Electra application which are not covered in the other papers. The particular application to the Electra provided new requirements which had not been previously encountered and which resulted in special equipment designs. The design methods described below resulted in enabling the advantages of static excitation and magnetic amplifier controls and protection to be successfully achieved.

System Description

A brief description of the system is given here in order that the relationship of the components can be understood in the discussions of their design. The system consists of a statically excited generator driven directly from the turboprop engine or a gear box mounted on the engine. This generator is air-cooled. Mounted in the generator housing is a permanent magnet generator which is used as a source of control and protective power independent of the main generator loads. The main generator consists of 3-phase windings on the stator and an 8-pole field rotor fed through slip rings. The output power from the stator is fed through the static exciter unit which is mounted in the main fuselage. The static exciter uses a voltage and current feed from the generator to supply power for excitation to the generator rotor field. Mounted near the static exciter is the voltage regulator which controls the amount of excitation power. Mounted adjacent to the voltage regulator is the control unit which supplies the control and protective functions and the necessary signals to the transfer relays.

Design Problems

The system had to meet special requirements for this particular application. The generator is electromagnetically rated at 60 kva; that is, double load is 120 kva. Lockheed Aircraft Corporation specified a generating system capable of delivering 90 kva (150% rated load) during flight from sea level to 30,000 feet when supplied cooling air per Fig. 1. The system is required to meet all load requirements at 5,700 rpm, corresponding to 380 cycles. This differs from the usual requirement for overloads at 400 cycles. The system is required to maintain 117 volts line to neutral at the end of 60 feet of feeders having a voltage drop of 3.54 volts at 150% load. The 60 feet of feeders are the connection from the outboard engine generators to the static exciters which are located in the fuselage. The voltage regulation sensing point is at the output of the exciters, which is close to the load bus. Under all environmental and operating conditions including 200% load the voltage is required to be maintained within ±4%. The voltage recovery limit requirements are shown in Fig. 2. The system is required to be capable of supplying under all operating conditions at the end of the feeders at least three per unit, 3-phase short-circuit current and at least five per unit single-phase-to-neutral short-circuit current. The control system is required to prevent power of the improper characteristics from being applied to any phase of the load equipment. This requirement is met by the use of overvoltage, undervoltage, underfrequency, and overfrequency protective devices. In addition, the system is required to protect the aircraft against damage due to short circuits on the main feeders to the aircraft structure. The control system is required to actuate the transfer relays properly so as to provide for complete automatic operation of the system. The components are required to be designed for maximum reliability and air-line serviceability and for 2,000 hours' life between overhauls. The following sections show what design features are incorporated in the components of the system in order to achieve these requirements and the results of these design features in meeting them, are presented.

Nomenclature

$A = NrAc$
A_c = area of one rotor conductor, square inches
$a = \dfrac{1}{1 + K_s/K_r}$
$b = a/2(K_I/K_r)$
I_r = rotor current, amperes
K = constant relating N to rotor current
$K_I = KK_I'$
K_I' = constant relating iron weight to N
K_r = constant relating rotor watts to N
$K_s = \dfrac{K_s'}{K^2}$
K_s' = constant relating stator conductor weight to N
l = height of iron of rotor, inches
N = stator turns in series per phase
N_r = turns per pole in rotor
P = number of poles
R = rotor resistance, ohms
X_d = synchronous reactance, ohms
Z_0 = zero sequence impedance, ohms
Z_2 = negative sequence impedance, ohms
γ = weight density of rotor conductor material number pounds per cubic inch
ρ = resistivity of rotor conductor material $\left(\dfrac{\text{ohms} - \text{inches}^2}{\text{inch}}\right)$
ϕ_{FL} = the full-load flux kiloline
W_{as} = width of stator slot, inches
W_{ac} = width of stator conductor, inches
K_m = tooth flux factor
K_ϕ = yoke flux factor
H_{as} = height of stator slot, inches
T_c = turns per coil

Paper 57-471, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 1, 1957; made available for printing May 16, 1957.

J. R. M. ALGER, W. E. WARDEN, W. O. HANSEN, and L. KLEIN are with the General Electric Company, Mr. Alger in Erie, Pa., Messrs. Warden and Hansen in Waynesboro, Va., and Mr. Klein in Schenectady, N. Y.

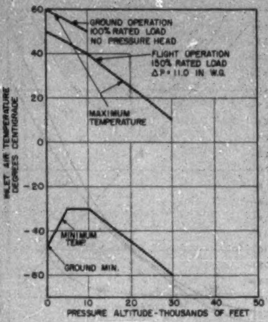


Fig. 1. Generator cooling air chart

H_{ac} = height of stator conductor, inches
A_{ac} = area of armature conductor, square mils
L_s = length of stator stack, inches
L_{cs} = length of wire in stator
Ω_{310} = resistance at operating temperature line to neutral (stator)
R_{310} = per-unit resistance at operating temperature (stator)
W = the desired stator loss, watts
W_s = computed stator loss
W_{cs} = weight of copper in stator
W_{aso} = width of armature slot opening at air gap
W_{is} = weight of iron in stator
E_{ia} = assumed per-unit induced voltage
L_{urp} = length of underside of pole face (one side)
K_{1c} = calculated value of K_1
D_{is} = final inside diameter of stator
K_{ad} = direct axis reactance constant
K_{aq} = the quadrative axis reactance constant
X_{ad} = the direct axis reactance per unit
X_{aq} = the quadrative axis reactance, per unit
X_a = the armature leakage reactance, per unit
E_{tr} = component of induced voltage in phase with current
E_{tx} = component of induced voltage in quadrative with current
E_{ic} = the induced voltage per unit as calculated

Generator Design

MINIMUM SPEED

The standard generators are often referred to as 380- to 420-cps (cycles per second) generators. However, the specifications normally only require overload testing at 400 cps. This application required overload operation at 380 cps (5,700 rpm) rather than the standard machine requirement for 400 cps (6,000 rpm). The decrease in rpm requires greater field current due to the decrease in induced voltage. The increase in field current is more than proportional to the speed ratio as the machine is operating under saturation at overloads. The increased field current will cause increased heating of the rotor which increases the rotor resistance. The increase in rotor resistance requires more field power and further increases the heating causing a cumulative effect. For example, a standard machine operating at 6,000 rpm and 1.5 load (90 kva) requires a field power at stabilized conditions of approximately 3 kw. To operate that machine at 5,700 rpm, 1.5 load, requires approximately 4 kw. Thus, for a 5% decrease in speed the rotor power requirements increase approximately 30%. This will overheat the rotor and also require a larger exciter. To keep the field power at 3 kw at 5,700 rpm, 1.5 load required the addition of 2 pounds of electromagnetic weight.

OVERLOAD

The voltage setting of this system is 117 volts at the end of the feeders. A ±4% tolerance is allowed on the voltage for all combinations of environmental and load conditions giving a minimum voltage of 112.2 volts. At double load the feeder voltage drop will be 4.7 volts. Therefore, the generator is designed to supply the power for the double load plus the feeder loss after the system has reached the maximum stabilized temperature for 1.5 load at 5,700 rpm. This requires power considerably in excess of the standard requirement in effect at the time the Electra design was initiated where the double-load voltage was allowed to droop without supplying any feeder power and at 6,000 rpm. This also required an increase in electromagnetic weight. Tests conducted on the Electra generator at sea level altitude, 5,700 rpm, and 57 C (degrees centigrade) inlet cooling air with an 11-inch water pressure drop across the generator showed the output voltage at the end of the feeder to be 113 volts under double load after the system had stabilized at 150% load. In the data for this test, feeder voltage drop in part was accounted for analytically, since the actual feeders were not yet installed in the test area.

THERMAL

Due to the requirement for continuous operation at 5,700 rpm, 150% load, the thermal conditions for this application are actually more severe than the standard class *C* conditions. As an example, a test was made at the most severe class *C* condition of stabilized full load, 6,000 rpm at 65,000 feet altitude with −12 C inlet cooling air at 11 inches water pressure drop across the generator. Under this class *C* condition the generator operated at cooler temperatures than for the most severe Electra conditions. The thermal conditions required the application of an advanced insulation system.

The insulation system developed for this generator consists of inorganic materials all generally acceptable within AIEE class *H*. The system is the type used in military class *C* machines.

The materials used are: asbestos; glass laminates, silicone binder; glass-mica-glass laminates, silicone binder; silicone rubber; silicone varnish.

The most vulnerable component, temperature-wise, of the insulation system is the varnish. Insulation system studies using the AIEE motorette test code as a guide are presently underway.

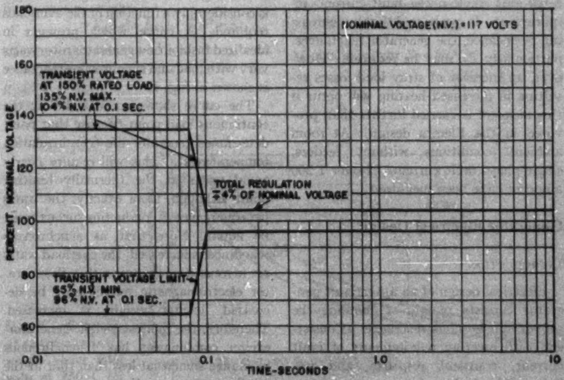


Fig. 2. Voltage recovery limits

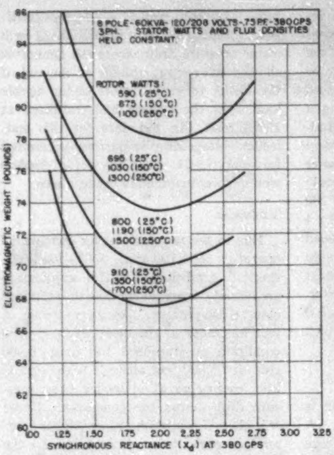


Fig. 3 (left). Electromagnetic weight versus synchronous reactance for various magnitudes of rotor watts

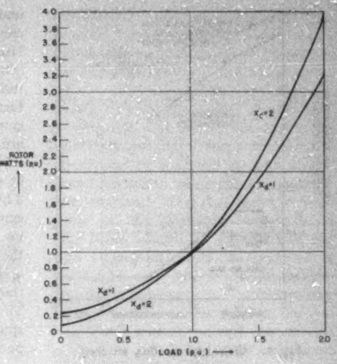


Fig. 4 (right). Rotor watts versus load

FAULT CURRENTS

All generator reactances are proportional to the square of the series turns per phase (N). Each reactance may be considered an inductance, proportional to N^2 times the permeance. Therefore, an important factor in optimizing system weight is to insure that the choice of X_d, which determines N, does not jeopardize the values of Z_0 and Z_2 needed to meet fault current or unbalanced load requirements. As a result of these factors, a small weight increase in the generator was accepted in order to keep the series turns per phase (N) at a value consistent with values of Z_0 and Z_2 that would meet the 5-per-unit single-phase fault-current requirement. It can be said that by increasing permeance, the generator reactances, particularly Z_2, may be reduced. However, an increase of stray load losses resulting in increased heating will occur if this measure is carried further than provided in the Electra design. At room ambient conditions without feeders, single-phase fault currents of over 1,200 amperes have been measured.

Generator Electrical Design

INTRODUCTION

Electrical design of an a-c aircraft generator consists mainly of choosing the proper synchronous reactance. Consistent with meeting requirements of fault current, transient response, and unbalanced load, the synchronous reactance will be chosen at the value to give minimum electromagnetic weight. For comparable designs having the same losses and saturation level, the generator electromagnetic weight varies with synchronous reactance X_d as in Fig. 3. That is, there is an optimum value below and above which the electromagnetic weight increases. For low synchronous reactance, the weight is primarily in iron while at high X_d a larger proportion of the weight is that of the conductors. A further consideration is the electromagnetic weight of the exciter. The exciter saturation level, and to some extent, exciter conductor weight, is determined by the overload excitation watts required. Therefore, exciter electromagnetic weight is considerably a function of the overloads required. A curve which presents in idealized fashion how generator rotor watts vary with load and synchronous reactance is shown in Fig. 4.

The curve shows equal watts at the continuous load point because that point determines generator life (i.e., insulation temperatures). This will require nearly equal watts at the thermally limiting point although, more exactly, the watts per square inch of conducting surface must be equal. Note that as synchronous reactance is increased, the overload watts are increased, so that even though generator electromagnetic weight may be decreased, exciter weight is increased. Therefore, the lightest weight generator-exciter combination has a synchronous reactance somewhat less than that at the minimum point on the curve of generator electromagnetic weight versus X_d.

The discussion following illustrates in simplified but ready terms why generator electromagnetic weight varies with synchronous reactance and how the digital computer is used to provide basic data to design lightweight a-c aircraft generators.

WEIGHT VERSUS SYNCHRONOUS REACTANCE

The following analysis applies to designs holding a fixed speed, a given number of poles, and the same design voltage.

The electromagnetics of an aircraft generator consist of conductor and iron weight. Minimum electromagnetic weight can be obtained for a-c generators by finding the proper weight ratio between conductor and iron. This minimum weight ratio can be described by the value of synchronous reactance, as the simplified analysis to follow will show.

The stator conductors are wound in slots with a certain number of turns (called N) in series per phase. The weight of those turns can be described by the equation:

Stator conductor weight $= K_s'N^2$

where K_s' is a constant. This is so because, if the number of turns in series per phase is doubled, which approximately doubles the weight, the wire area also must double, which also approximately doubles the weight, in order to maintain the same copper loss as before doubling N. In order to compare designs for weight fairly, copper losses are maintained the same. Of course, this is simplified because actually what must be done is to compare designs for weight that cool to the same temperature (i.e., have constant watts per square inch and comparable coolant velocities).

Now, how does rotor conductor weight

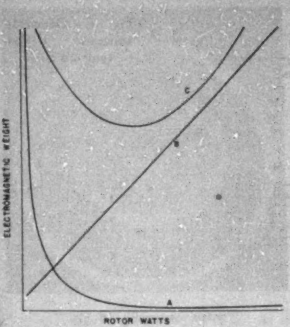


Fig. 5. System electromagnetic weight versus rotor watts

A—Generator
B—Exciter
C—Sum of generator and exciter

vary with N to maintain constant copper loss? As N is varied, the rotor current as shown in the following varies approximately as $1+KN$ for a 0.75 power factor design where K is a constant. For constant terminal voltage (strictly true only for rotating exciter machines) the full-load rotor ampere-turns (FLAT) are nearly the sum of ampere-turns required at no-load rated voltage and the ampere-turns required to supply rated current to a 3-phase short circuit.

By holding constant air-gap flux density, which is usual practice, the no-load rotor current is relatively independent of N. However, as N is varied, the ampere-turns on 3-phase short circuit vary likewise for constant output current. Therefore, in per-unit turns, the rated-load rotor ampere-turns vary as $1+KN$:

$$\text{FLAT}\ \alpha(1+KN)$$

Rotor watts vary as rotor current squared as explained in the following and so rotor wire weight varies as $(1+KN)^2$ in order to maintain constant rotor copper loss. Rotor watts may be written as:

$$P_r=I_r^2R=\left(\frac{\text{FLAT}}{N_r}\right)^2\left(\frac{\rho lPN_r}{A_c}\right)$$
$$=(\text{FLAT})^2\left(\frac{\rho lP}{A}\right)$$

where

$$A=N_rA_c$$

It is seen that rotor power is inversely proportional to "window" wire area and independent of the number of rotor turns (nelgecting insulation space). Then, since wire weight is

$$\text{Rotor conductor weight}=\gamma lA_cN_rP=\gamma lAP$$

Therefore,

$$A=\frac{\text{weight}}{\gamma lP}$$

by substitution:

$$P_r=(\text{FLAT})^2\frac{\gamma\rho(lP)^2}{\text{weight}}$$

or for constant rotor watts:

$$\text{Rotor conductor weight}=K_r(1+KN)^2$$

where K_r is a constant, assuming relatively small length changes.

Finally, the sum of stator and rotor iron weight will vary inversely as N because comparable designs must have the same design flux densities, or otherwise saturation differences will obscure the weight comparison:

$$\text{Iron weight}=K_I'/N$$

where K_I' is another constant.

Putting the weight variations together:

$$\text{Electromagnetic weight}$$
$$=K_s'N^2+K_r(1+KN)^2+\frac{K_I'}{N}$$

Next, it is desired to express weight variation in terms of synchronous reactance. By definition:

$$X_d=\frac{\begin{array}{c}\text{Rotor amperes for rated current}\\ \text{3-phase short circuit}\end{array}}{\begin{array}{c}\text{Rotor amperes on air-gap line}\\ \text{for rated voltage}\end{array}}$$

and so, neglecting saturation at rated load, $X_d=KN$, because, as noted previously, rotor amperes for no-load rated voltage are relatively independent of N while the rated current short-circuit rotor amperes increase directly as N. Then:

$$\text{Electromotive weight}$$
$$=K_sX_d^2+K_r(1+X_d)^2+K_I/X_d$$

where

$$K_s=\frac{K_s'}{K^2}$$

Table I. Generator Performance Versus Air Gap at 1 1/2 Load, 5,700 Rpm, 20,000 Feet, 10 C Air

Air-Gap Length	Rotor Amperes	Rotor Watts	Rotor Rise, C
0.025	49.1	2,670	256
0.030	50.8	2,660	222

and

$$K_I=KK_I'$$

This equation shows that as X_d is raised, the iron weight decreases and the conductor weight increases. Clearly an optimum X_d for minimum weight can be obtained because a point is reached beyond which conductor weight will increase faster than iron weight is decreased. The minimum weight X_d is found where the first derivation of electromagnetic weights with respect to X_d is zero.

This is, then, the real positive root solution to the following equation (see the appendix):

$$X_d^2(X_d+a)=b$$

where

$$a=\frac{1}{1+K_s/K_r}$$

$$b=\frac{a}{2}\left(\frac{K_I}{K_r}\right)$$

Experience shows that the values of a are small compared to b. Then the solution is

$$X_d=b^{1/3}\left(1-\frac{a}{3b^{2/3}}\right)$$

The term in parentheses is usually about 0.01 to 0.03. Neglecting it:

$$\text{Minimum weight } X_d=b^{1/3}$$
$$=\left[\frac{K_I}{2(K_s+K_r)}\right]^{1/3}$$

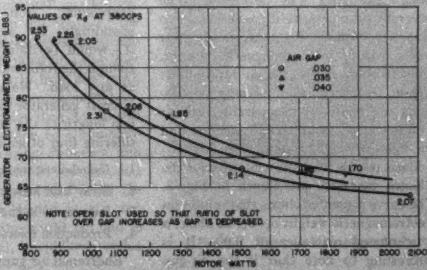


Fig. 6. Generator electromagnetic weight versus rotor watts at 250 C for various air gap lengths

The values for the Electra generator are

$X_d = 2.0$ per unit
Stator copper weight = 5.8 pounds
Rotor copper weight = 8.1 pounds
Iron weight = 49.0 pounds
Amortisseur = 2.0 pounds

For these values, the calculated constants are

$K_s = 1.45$
$K_r = 0.90$
$K_l = 98$

Putting these back into the electromagnetic weight equation, and letting X_d be an independent variable:

Electromotive weight

$$=1.45X_d^2+0.9(1+X_d)^2+\frac{98}{X_d}$$

$$=2.35X_d^2+1.8X_d+\frac{98}{X_d}$$

Where this is plotted as Fig. 3, then the minimum weight X_d is

$$X_d=\left(\frac{98}{4.70}\right)^{1/3}=2.75$$

If saturation were considered, a lower value of X_d would have resulted due to the extra iron weight necessary to maintain acceptable flux densities. Also, the static exciter and stator leakage impedance require additional voltage that again adds to the iron requirements in order that rotor power be maintained at a reasonable value. Only a complete design method can include all the factors that affect the electromagnetic weight. But, from this rough derivation, the conclusion may be drawn that the ratio of iron to copper weight determines the total electromagnetic weight; and, giving due consideration to system requirements, a minimum weight may be achieved.

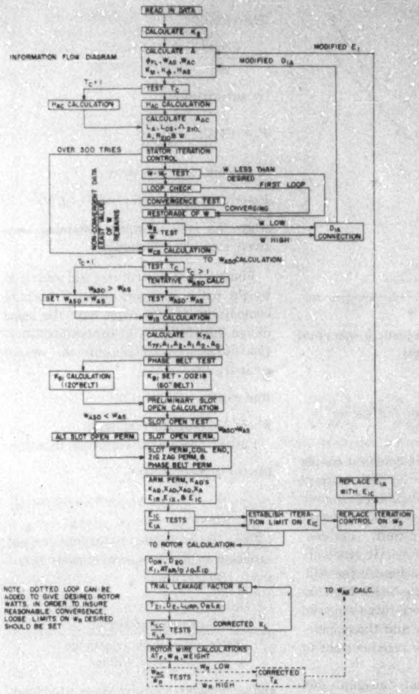

INFORMATION FLOW DIAGRAM
READ IN DATA
CALCULATE K_g
CALCULATE A
$\phi_{FL}, W_{AS}, W_{AC}$
K_M, K_ϕ, H_{AS}
MODIFIED E_1
MODIFIED D_{1A}
$T_C = 1$
TEST T_C
H_{AC} CALCULATION
H_{AC} CALCULATION
CALCULATE A_{AC}
$L_A, L_{CS}, \Omega_{210}$,
A, R_{210} & W
OVER 300 TRIES
STATOR ITERATION CONTROL
W-W_L TEST
W LESS THAN DESIRED
LOOP CHECK
FIRST LOOP
CONVERGENCE TEST
CONVERGING
NON-CONVERGENT DATA LEAST VALUE OF W REMAINS
RESTORAGE OF W
$\frac{W_S}{W}$ TEST
W LOW
W HIGH
D_{1A} CONNECTION
W_{CS} CALCULATION
TO W_{ASO} CALCULATION
$T_C = 1$
TEST T_C
$T_C > 1$
TENTATIVE W_{ASO} CALC
$W_{ASO} > W_{AS}$
SET $W_{ASO} = W_{AS}$
TEST W_{ASO}-W_{AS}
W_{IS} CALCULATION
CALCULATE K_{TA}
$K_{TT}, A_1, A_2, A_1 A_2, A_0$
PHASE BELT TEST
K_{B1} CALCULATION (120° BELT)
K_{B1} SET = .00218 (60° BELT)
PRELIMINARY SLOT OPEN CALCULATION
$W_{ASO} < W_{AS}$
SLOT OPEN TEST
$W_{ASO} = W_{AS}$
ALT SLOT OPEN PERM
SLOT OPEN PERM
SLOT PERM, COIL END, ZIG ZAG PERM, & PHASE BELT PERM
ARM PERM, K_{AD}'S
$K_{AQ}, X_{AD}, X_{AQ}, X_A$
E_{1R}, E_{1X}, & E_{1C}
REPLACE E_{1A} WITH E_{1C}
$\frac{E_{1C}}{E_{1A}}$ TESTS
ESTABLISH ITERATION LIMIT ON E_{1C}
REPLACE ITERATION CONTROL ON W_S
TO ROTOR CALCULATION
D_{OR}, D_{20}
$K_1, AT_{AR}, I_D, J_Q, E_{1D}$
TRIAL LEAKAGE FACTOR K_L
$T_{21}, D_Z, L_{URR}, O_R, L_R$
$\frac{K_{LC}}{K_{LA}}$ TESTS
CORRECTED K_L
TO W_{AS} CALC.
ROTOR WIRE CALCULATIONS
AT_F, W_R, WEIGHT
$\frac{W_{RC}}{W_R}$ TESTS
W_R LOW
W_R HIGH
CORRECTED T_A
NOTE: DOTTED LOOP CAN BE ADDED TO GIVE DESIRED ROTOR WATTS. IN ORDER TO INSURE REASONABLE CONVERGENCE, LOOSE LIMITS ON W_R DESIRED SHOULD BE SET

Fig. 7. Generator computer design synthesis deck

Digital Computer Design Synthesis

Program Results

As an interim step in the design synthesis of lightest weight aircraft electric systems, an International Business Machines *650* computer program for generator design synthesis was developed. The program provides the data necessary to plot synchronous reactance and rotor watts versus electromagnetic weight. This points out the design for lightest weight generator; and, because the designs are relatively comparable, exciter weight can be computed and added to find the lightest weight system. See Fig. 5.

Basic results from the computer for the Electra generator are shown in Fig. 3. The curve presented shows that minimum electromagnetic weight occurs at about a synchronous reactance of 2.00. It is interesting to note that if the air-gap length is varied assuming constant fringing, the synchronous reactance will vary inversely. For a small air gap very high X_d will result. However, a given design will have less rotor copper loss at all loads if the air gap is shorter, although X_d is higher. Fig. 6 shows a calculated weight and X_d variation with air gap. It is easy to see that the lightest weight design would have zero length air gap. In that case, the rotor would not have to supply air-gap energy and could devote it all to overcoming iron saturation. Correspondingly, for a zero-length air gap, X_d would be infinite, by definition, although the rotor watts would be the least at any load. Naturally, therefore, the aircraft generator designer holds his air-gap length to the smallest possible, considering four other items:

1. Unbalanced magnetic pull.
2. Stray load losses.
3. Harmonics in the voltage.
4. Production tolerances.

Shortened air gaps require increasingly close production tolerances in order to avoid gap eccentricity. The minimum allowable gap from that standpoint is the one that lightweight mechanical construction will just still hold in place and with which the factory organization can maintain reasonable concentricity. Furthermore, shortened air gaps make relatively more difficult the task of holding down flux pulsations across the pole faces due to poorer fringing. These flux pulsations give rise to extra losses in the iron surfaces and cause extra rotor heating. It is interesting to note that because of this condition a given design will have an "optimum" air gap for minimum rotor watts, as the data in Table I illustrate.

For smaller air gaps, iron heating will cause more rotor losses than the reduced rotor current can overcome. At larger air gaps the increasing rotor current increases the losses. Of course, this is based on an existing design where fringing cannot be improved while decreasing the gap.

There are many factors to investigate

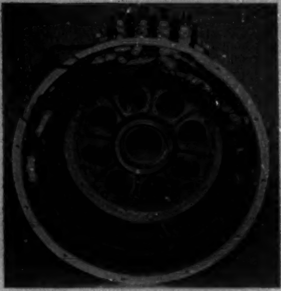


Fig. 8. Generator stator

whose impact on weight are rather obscure. Some of these are

1. Diameter-to-length ratio.
2. Pitch and distribution.
3. Air-gap length.
4. Coil throw.
5. Harmonic content.
6. Insulation thickness.
7. Saturation level.
8. Conductor material.

The computer study was used to optimize items 1 through 4.

Program Development

A flow chart of the computer design synthesis deck is given as Fig. 7. The stator is designed first (Fig. 8), requiring stator copper loss to be held to an input figure, (within ±1%). Stator iron design maintains specified flux densities for the given stator turns. Finally, the rotor is designed holding specified flux densities. The correct leakage impedances are fed back in both rotor and stator to give the flux densities correctly.

The beauty of synthesis is that design flux densities are specified as inputs. Therefore, only one point on the magnetization curve need be put into the deck for each design density. The tapering stator tooth is taken care of by an empirical relationship derived to fit calculated magnetization curves for various tip-to-slot-width ratios including slot flux.

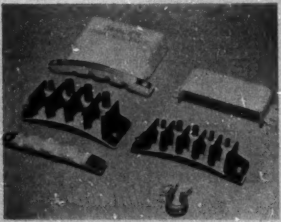


Fig. 9. Generator terminal blocks

One curious effect which was not at first anticipated arose on early trials of the deck. In order to design a stator with the desired stator watts (within ±1%), the stator bore is varied. For example, if stator watts calculated on the first pass are too low, inside diameter is increased. Outside diameter happens to be a piece of input data and so the result is to "squeeze" the slots and so increase the stator watts. Naturally, convergence on the desired watts is assured, because inside diameter is increased until stator watts reach the specified value. However, if stator watts are too high on the first pass, it is possible that on the next pass stator watts will be even higher, even though inside diameter is decreased. This is because as stator bore is decreased, to allow a deeper slot and so lower stator watts, the slot becomes narrower so that a point is reached beyond which slot area becomes *less* rather than greater and so stator watts increase rather than decrease. If this case arises, it simply means that the desired stator watts are too low to be designed into the specified package. To circumvent this situation, the deck compares the newly calculated watts with the previous value for the case of stator watts greater than desired. If the latest value is lower than the previous, calculations continue. If not, the deck accepts the "lowest watts" design and proceeds. The design study was based on the use of rectangular stator slots.

Generator Mechanical Design

MECHANICAL WARNING DEVICE

Recognition and warning of bearing failures in aircraft generators is a safety feature desired by many customers. To accomplish this, a mechanical warning device is incorporated in one tooth of the stator core. This device consists simply of a copper strip embedded in a slot liner which insulates the strip from the tooth iron. Fig. 9 shows the component details. Bearing failure will cause the rotor to pull over onto the stator due to a combination of increased bearing clearance and magnetic unbalanced force. A test on a 15-kva generator showed that with the mechanical warning device in the top tooth (so that gravity would not aid in activating the device), the pullover action of the rotor was ample to smear the copper strip onto the tooth iron and thereby close the circuit to the warning light in the cockpit.

Fig. 10. Generator rotor

QUICKLY REMOVABLE BLAST CAP

Brush inspection and maintenance are enhanced by incorporation of a simply removable inlet cap. The problem in this case is safety wiring. By using a clamping band design (which springs the band down onto the cap) only one safety wire need be removed in order to release the cap.

TERMINAL BLOCK DESIGN

Some desirable terminal block design features incorporated in the Electra generator are

1. Barrier to prevent short circuits of cable terminals.
2. Covers to prevent short circuits from extraneous matter.
3. Current conduction from equipment to cable terminals by surface-to-surface contact and not through screw threads.
4. Removal of mechanical loads from cable terminals.

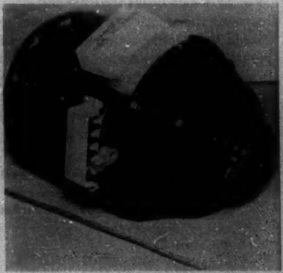


Fig. 11. External view of generator

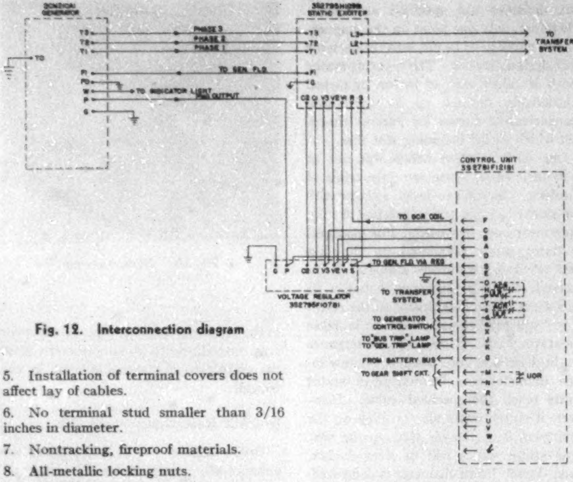

GENERATOR
T3
T2
T1
T0
F1
PD
W
P
G
PHASE 3
PHASE 2
PHASE 1
TO GEN. FLD.
TO INDICATOR LIGHT
PMG OUTPUT
STATIC EXCITER
L3
L2
L1
TO TRANSFER SYSTEM
C2 C1 V3 V2 V1 R S
CONTROL UNIT
TO GCR COIL
TO GEN. FLD. VIA REG
G P C2 C1 V3 V2 V1 S
VOLTAGE REGULATOR
TO TRANSFER SYSTEM
TO GENERATOR CONTROL SWITCH
TO "BUS TRIP" LAMP
TO "GEN. TRIP" LAMP
FROM BATTERY BUS
TO GEAR SHIFT CKT.
F
C
B
A
D
S
E
O
H
P
Y
J
K
G
M
N
R
L
T
U
V
W
ACR
UCR
UOR

Fig. 12. Interconnection diagram

5. Installation of terminal covers does not affect lay of cables.

6. No terminal stud smaller than 3/16 inches in diameter.

7. Nontracking, fireproof materials.

8. All-metallic locking nuts.

9. Terminal number and color coding.

These features are incorporated into the generator design as shown in Fig. 10. Each cable is separated by barriers from its neighbors. Current is carried from the stator through flag-type connectors against which the cable terminals are placed in direct contact. Further, the main power terminal block cover is designed with an integral cable clamp assembly which binds the main power cables firmly in place as the cover is tightened. This releases the terminal studs and cable terminals of mechanical forces. The auxiliary terminal block has a simple loop clamp installed on the generator directly in front of the block, through which the cables are laid and clamped. Once again mechanical forces are stopped at the clamp.

FAN COOLING

Fig. 11, which shows the rotor assembly, illustrates the wedge fan that is used to provide the cooling required for ground operation of the generator. The wedges are extended on the exhaust end of the generator and provide centrifugal fan action. Typical temperatures with and without this fan are shown in Table II.

Table II. Fan Action CM210 at 6,000 Rpm

Item	With Fan	Without Fan
Rotor temperature rise at full load	143	187
Air rise at full load	55	99
Approximate air flow, cubic feet per minute	200	100

METHOD OF BALANCE

Fig. 11 also illustrates the end construction of the rotor. High-strength bolts are used to hold down small caps which are placed over the end turns. These caps are made of titanium in the initial construction, and in order to facilitate balance, steel caps of comparable yield and tensile strength are substituted where needed. The weight difference between titanium and steel therefore provides a nice way of obtaining balance.

PHYSICAL CHARACTERISTICS

The use of static excitation has many advantages in the design of the generator. It eliminates the designing of the rotating exciter with its commutator. Because of the elimination of the rotating exciter the cooling problems are reduced. The generator is smaller, lighter, and has less overhung moment. The Electra generator is 11 inches outside diameter and $13^7/_8$ inches long including blast cap, weighs 91.5 pounds, and has an overhung moment of approximately 435 inch-pounds. The reduced length of the static exciter generator helps to solve the installation problem in the accessory section of the Electra nacelle.

Static Exciter and Regulator

This paper will be limited to discussing those aspects of static exciter and regulator design and performance peculiar to the Lockheed Electra. The interconnection diagram of the system is shown in Fig. 12.

STATIC EXCITER

Briefly stated, the exciter supplies field excitation current by feeding back to the generator, a rectified power proportional

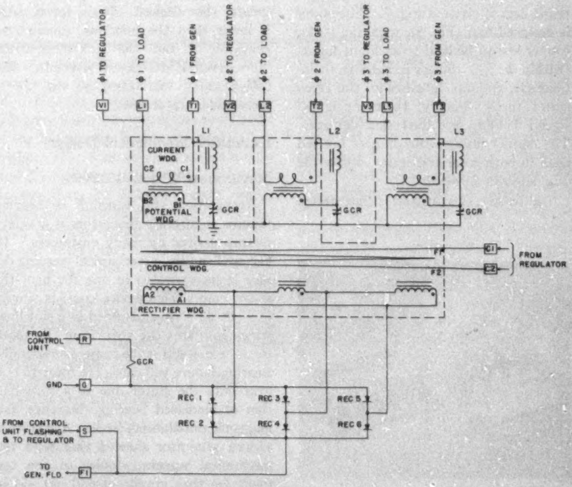

ϕ1 TO REGULATOR
ϕ1 TO LOAD
ϕ1 FROM GEN
ϕ2 TO REGULATOR
ϕ2 TO LOAD
ϕ2 FROM GEN
ϕ3 TO REGULATOR
ϕ3 TO LOAD
ϕ3 FROM GEN
V1
L1
T1
V2
L2
T2
V3
L3
T3
L1
L2
L3
CURRENT WDG
C2
C1
B2
B1
POTENTIAL WDG
GCR
GCR
GCR
F1
F2
C1
C2
FROM REGULATOR
CONTROL WDG
A2
A1
RECTIFIER WDG
FROM CONTROL UNIT
R
GCR
GND
G
REC 1
REC 2
REC 3
REC 4
REC 5
REC 6
FROM CONTROL UNIT FLASHING & TO REGULATOR
S
TO GEN. FLD
F1

Fig. 13. Static exciter diagram

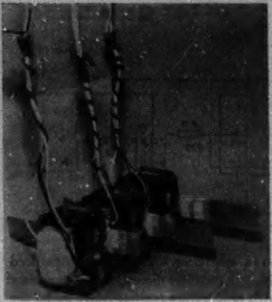


Fig. 14. Saturable current potential transformer

to a combination of line voltages and currents. The current and voltage feeds are combined vectorially in accordance with the approximate relationship:

$$If = K\mathbf{E}_t + j\mathbf{I}_l Xd \qquad (1)$$

where

If = field current in d-c amperes
K = constant of proportionality
$\mathbf{E}_t$ = generated voltage vector
$\mathbf{I}_l$ = line current vector
Xd = generator synchronous reactance

Note that $\mathbf{E}_t$ is also a function of Xd.

The voltage and current signals are so proportioned that, for normal loads, the static exciter will supply correct generator excitation to maintain almost constant voltage, without any action from the regulator.

Essentially, the exciter consists of the following parts (see diagram, Fig. 13):

1. Saturable current potential transformer (SCPT), in which the voltage and current feeds are combined vectorially and into which a control signal from the regulator is introduced.
2. Three-phase bridge rectifier which rectifies the exciter output.
3. Linear reactors for shifting the voltage signal in quadrature with the current signal.
4. Generator control relay (GCR) which, when de-energized, short-circuits the SCPT potential windings to de-energize the system. The GCR is controlled by the cockpit generator control switch (GCS) and the control unit as explained in the "Control Unit."

The SCPT is a 3-phase unit employing *C* core magnetic material and is shown in Fig. 14. A continuous regulator control winding is wound around all three phases, the other windings (potential, current, and output) are all wound individually around each core. Use of the *C* cores with their continuous grain orientation makes it possible to use less magnetic material than would be needed for cores made of punched material. A continuous control winding is used to eliminate the problems associated with high voltages induced in individual control windings.

The GCR relay which is used to shut the system down has a relatively easy function by short-circuiting the windings of the exciter. The closing of the GCR contacts kills the excitation to the generator and the system immediately starts building down. Thus, the relay which is used to turn the system off does not have to open any high power conditions. When the static exciter system is shut down, its output voltage is less than one volt and thus, energy is not available to feed a fault. In the rotating exciter systems where the exciter field is opened in order to shut down, a voltage from 50 to 150 volts is available to feed the fault.

RATING

The requirement of 90-kva continuous operation at 5,700 rpm affects the exciter design from the standpoint of heating and output power requirement.

The consideration of maximum temperature in the exciter determines primarily the conductor size of the winding on the SCPT and linear reactor. Tests have been run at 90 kva, 380 cps, and the exciter at +71 C to determine the maximum temperatures of various parts of the exciter. Some of the temperatures reached during a 5-hour run at these conditions were as follows:

SCPT core: +157 C (stabilized after 1$^1/_4$ hours of operation)
SCPT current winding: +170 C (stabilized after 1$^1/_4$ hours of operation)

The SCPT current winding is a copper strap having a cross-sectional area of 62,500 square mils.

It is interesting to note that the exciter linear reactors reach their maximum temperature while the system is at no load. During tests at +71 C and 380 cps, a reactor core stabilized temperature of +204 C was obtained.

When load current flows through the current coil of the SCPT at high lagging power factors, it induces a voltage in the SCPT winding supplied by the linear reactor. This induced voltage reduces the current through the linear reactor.

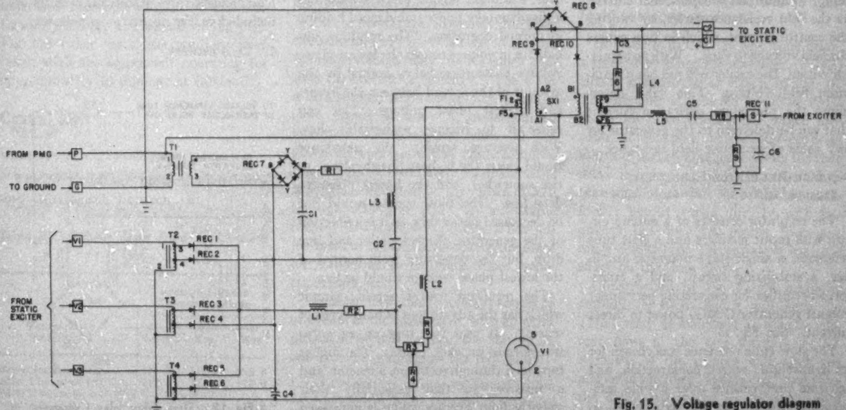


Fig. 15. Voltage regulator diagram

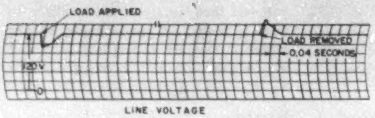


Fig. 16 (above). Transient voltage response

Full load (60 kva at 0.75 pf) applied and removed frequency = 380 cps

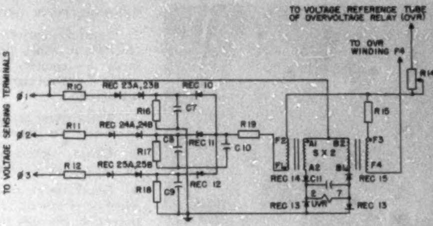


Fig. 17 (right). Undervoltage relay diagram

The static exciter is a highly efficient means of conversion. Following is a tabulation of the exciter losses and efficiency while operating at 1.5 loads. These figures are based on 380-cps operation with generator field resistance = 1.2 ohms, 50-ampere field current, exciter components at their maximum temperature with +71 C ambient, and an assumed 2-volt drop across the rectifiers. Loss figures are probably good to ±10% but the efficiency value is more accurate.

I^2R losses:	
Linear reactor =	16.8 watts
Output winding =	62.9 watts
Potential winding =	6.0 watts
Current winding =	32.0 watts
Control winding =	4.3 watts
Total I^2R losses =	122.0 watts
Core loss (SCPT) =	60.0 watts
Core loss (Linear Reactor) =	10.0 watts
Rectifier loss =	100.0 watts
Total losses	292.0 watts
Efficiency = 92.2%	

The matter of maximum design field voltage is a joint exciter-generator consideration. Essentially, the exciter and regulator combination form a constant current device. For any given generator load and frequency the regulator will attempt to maintain constant field current as the field resistance varies, by varying the control current, which in turn causes the field voltage to vary. With no regulator output, the exciter will put out a maximum field voltage. This value determines the maximum excitation current that can be delivered to the generator for any value of generator field resistance.

VOLTAGE REGULATOR ELECTRICAL DESIGN

The regulator consists of a sensing circuit with input rectifiers and a glow tube reference, a single-stage magnetic amplifier, a stabilizing circuit, and a transformer-rectifier for converting permanent magnet generator (PMG) power to direct current. See Fig. 15.

The glow tube reference was chosen for its lightweight, rugged construction, and accurate performance over a wide temperature range. Service experience on thousands of static regulators testifies as to its reliability.

Highest phase voltage sensing is a circuit design that eliminates a dangerous condition that can affect safety in flight. In most systems in use today, the voltage regulator senses the average of the three line voltages and regulates on this average. It is possible (and has occurred) for one or two phases to have an overvoltage with reduced voltage on the other phases. A faulted feeder or line in the distribution circuit can cause this unsafe condition with 3-phase average sensing. To see the reason for this, consider this example. *L1* is faulted at the bus and at zero voltage and *L2* and *L3* are still feeding power. The average voltage is low and the regulator increases excitation. Since overvoltage protection sees the same value as the regulator, it does not trip. The voltage on *L2* and *L3* can rise as high as 160 volts and cause damage to essential loads so that they would not be available even if transferred to another generator.

A solution to this problem is regulation of the highest phase voltage. This is accomplished by the addition of a capacitor in the regulator sensing circuit. However, a straight highest phase sensing has a disadvantage under unbalanced loading in normal operation. The resulting unbalance in voltage causes the lowest phase voltage to decrease below normal by the amount of the spread between the lowest and highest phase voltage. This can cause it to become undesirably low. With average sensing, the unbalance would cause the highest phase voltage to rise somewhat, and the lowest phase to drop less. The total spread would not be decreased noticeably, as it is a function of the generator characteristics and line drop, but the departure from normal of the lowest phase voltage would be less.

The regulator uses a sensing circuit which has the advantages of both types of sensing. It uses two taps on an autotransformer on each phase. The higher taps feed through rectifiers, a reactor, and a resistor, so that essentially their smoothed-out average value is impressed on a capacitor. The lower taps feed through rectifiers directly to the capacitor with peak values normally just under the voltage from the reactor and resistor. Thus, in normal operation the average phase voltage on the capacitor constitutes the signal to the reference circuit. However, if unbalance causes one phase voltage to rise enough so that its peak value from the lower tap exceeds the average value from the higher taps, this highest phase voltage will become the signal to the reference circuit. Thus, average sensing is used for normal operation, but highest phase sensing "takes over" when unbalance causes one phase voltage to rise to the limiting level.

The error signal from the reference circuit is amplified through the magnetic amplifier whose output in turn controls the saturable current-potential transformer in the exciter. Close voltage regulation is provided by the use of a high-gain magnetic amplifier in the voltage regulator. The system servo loop includes the time constants of the generator, exciter, and regulator. In order to provide stability to this loop and yet maintain fast transient voltage response to load changes three stabilizing circuits are included in the regulator as follows:

1. Lead network.

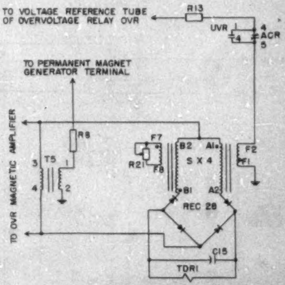

TO VOLTAGE REFERENCE TUBE OF OVERVOLTAGE RELAY OVR
R13
UVR
4
ACR
5
4
TO PERMANENT MAGNET GENERATOR TERMINAL
TO OVR MAGNETIC AMPLIFIER
R8
T5
3
4
1
2
F7
R21
F8
B2
A1
S X 4
F2
F1
B1
A2
REC 28
C15
TDR1

Fig. 18. Time delay relay 1 diagram

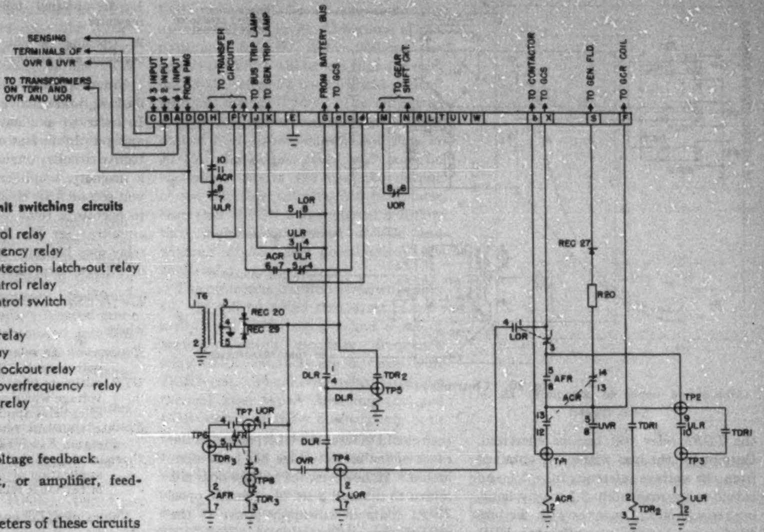


Fig. 19. Control unit switching circuits

ACR = auxiliary control relay
AFR = auxiliary frequency relay
DLR = differential protection latch-out relay
GCR = generator control relay
GCS = generator control switch
LOR = lockout relay
OVR = overvoltage relay
TDR = time delay relay
ULR = undervoltage lockout relay
UOR = under and overfrequency relay
UVR = undervoltage relay
TP = test point

2. Generator field voltage feedback.

3. Regulator output, or amplifier, feedback.

Values for the parameters of these circuits were obtained empirically.

A typical transient response trace showing full load at 0.75 power factor applied and removed at 380 cps is seen in Fig. 16.

Tests over the temperature range −55 C to +71 C indicate that the regulation up to 90 kva is well within the specified limits. Typical data taken with a set of prototype equipment are shown in Table III. These data were taken with a no-load line-to-neutral voltage setting of 116 volts at room temperature and the tests were run at 380 cps.

In the case of the +71 C test, the equipment had been operating at 90 kva continuously during the period of the test. For the other tests, the readings were taken with the equipment operating for approximately 10 minutes at full load.

Control Unit

UNDERVOLTAGE PROTECTION

The undervoltage protective circuitry must distinguish between two types of sustained undervoltages; those caused by excessive or fault current drawn by the load, and those caused by a fault or malfunction within the generator system. In each case, it must take proper action in order to prevent damage to the aircraft or its equipment. At the same time, it must be capable of ignoring transient undervoltages caused by sudden load application. Conventional differential protection is included in the Electra production models but was not included in the prototype systems.

Table III. Typical Data From Temperature Tests

Temperature Conditions	Average 3-Phase Voltage Readings		
	No Load	60 Kva	90 Kva
Room temperature	116.0	117.0	116.8
After 4 hours at +71 C	117.5	118.2	115.0
After 2 hours at −55 C at no load	117	118.0	117.8

The undervoltage magnetic amplifier circuit shown in Fig. 17 is used to operate the undervoltage relay (UVR) which controls the protective circuitry. The UVR is de-energized when the voltage on one or more phases is below a predetermined minimum and energized when all the phase voltages are above this minimum. The trip point, tentatively set at 105 volts ±2%, is determined by the values of *R10* through *R18* and *R15*; an adjustable resistor *R14* is used to provide a fine adjustment of the trip point and compensate for variations in reference tubes, relay trip points, etc.

Consider now the functional sequence of events in the case where, with all three phases at normal voltage, the voltage on one or more goes below the trip point:

1. If the voltage returns to normal within 3.0 seconds±25% (by removal of excessive load from the bus or tripping of load breakers), the UVR is re-energized and there is no action taken by the protective circuitry.
2. If the voltage does not return to normal within the 3 seconds, the bus will be disconnected from the generator.
3. If the voltage then returns to normal within an additional 1.0 second±50% the generator remains energized but the bus will be locked off the system.
4. If the voltage does not return to normal within the additional 1 second, indicating a feeder or generator fault, the generator will become de-energized and lock out of the system. The bus, however, will be transferred to another generator.

In order to obtain the first time delay, a time delay magnetic amplifier (*TDR1*) circuit is employed as shown in Fig. 18. The amplifier is powered from the PMG via a saturable transformer which provides a constant volt-second-per-cycle source of energy in order to make the time delay independent of PMG output fluctuations. Note that the saturable transformer output varies directly with frequency but over the usable frequency range 380–420 cps, this variation has little effect on the time delay.

Construction of the *TDR1* magnetic amplifier is such that it will keep the *TDR1* relay de-energized as long as current flows in the bias winding *F1-F2*. Removal of the current will start the magnetic amplifier "timing out" and after the delay determined by short-circuited winding *F6-F7* and resistor *R21* setting,

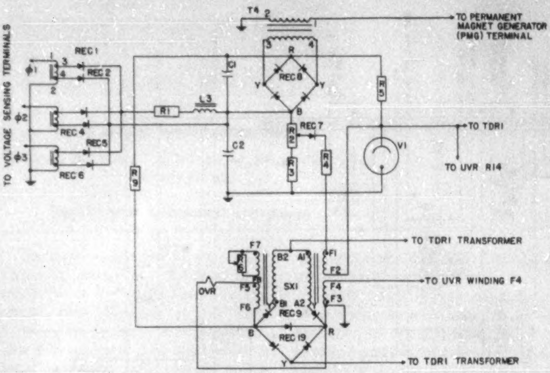

T4
2
1
TO PERMANENT MAGNET GENERATOR (PMG) TERMINAL
3
4
R
REC 8
Y
Y
B
TO VOLTAGE SENSING TERMINALS
φ1
φ2
φ3
REC 1
REC 2
REC 4
REC 5
REC 6
R1
L3
C1
C2
REC 7
R2
R3
R4
R5
R9
V1
TO TDR1
TO UVR R14
TO TDR1 TRANSFORMER
F7
R6
B2
A1
F1
F2
F5
SX1
F4
F3
OVR
F6
B1
A2
REC 9
REC 19
B
R
Y
TO UVR WINDING F4
TO TDR1 TRANSFORMER

Fig. 20. Overvoltage relay

the *TDR1* relay will become energized. Current for the bias winding is obtained from the voltage reference tube *V1* and is held to approximately 3.5 ma by limiting resistor *R13*. In series with the bias winding are a set of normally open UVR contacts and normally closed auxiliary control relay (ACR) contacts. The ACR is energized when the UVR picks up after field flashing and remains energized regardless of UVR operation, as long as the generator control switch (GCS) is closed and the frequency is within prescribed limits. This contact arrangement prevents removal of bias current and timing out when the generator is up to speed but the field has not been flashed, yet allows the relay to be energized when the voltage falls below the trip point after build up.

With reference to Fig. 19 which shows the unit's internal switching circuits, the sequence of operations after the dropping out of the UVR is as follows: The *TDR1* magnetic amplifier starts timing out. If the voltage returns to normal within the 3 seconds period, the UVR will be re-energized and the timing out action will be halted. If the voltage does not go back to normal, the *TDR1* relay will be energized. A set of *TDR1* contacts will then energize the undervoltage lockout relay (ULR), contacts of which operate the transfer circuitry to take the generator off the bus. The ULR is self-sealing in order to prevent recycling of the circuitry if the voltage returns to normal. In order to drop out the ULR, it is necessary to open the GCS and then close it.

If now the voltage returns to normal, there is no further operation of the circuitry. The GCS may be opened and then closed to reset the ULR, provided the cause of the undervoltage has been eliminated. If, however, the voltage does not return to normal, a set of normally open *TDR1* contacts will apply power to the coil of time delay relay *2* (*TDR2*), a thermal-type time-delay relay. In one additional second the *TDR2* contacts will close, energizing the differential lockout relay (DLR). A set of DLR contacts will energize the lockout relay (LOR) which will interrupt the coil current to the generator control relay in the static exciter and the bus transfer is reset. De-energization of this last relay will disable the exciter, thereby de-energizing the generator. Both the DLR and LOR have self-locking contacts, the former fed from the battery bus and the latter from the unit's TR via the GCS. Thus, to attempt to re-energize the generator it is necessary to open and close both the GCS and the connection between the unit and the battery bus.

OVERVOLTAGE PROTECTION

Sustained overvoltages can arise from several causes, including the following:

1. An internal failure of the voltage regulator.
2. Opening of the leads between the regulator and exciter.

Transient overvoltages arise from field flashing, and removal of loads and faults. In order to eliminate nuisance tripping and yet obtain fast operation of the protective circuitry on sustained overvoltages, a magnetic amplifier having a constant volt-second trip time delay is employed to operate a relay within the protective circuitry; see Fig. 20. The overvoltage relay uses highest phase takeover sensing of voltage. Expressed mathematically:

$$K=\int_0^T (E_T-UTP)dt$$

where

K = overvoltage relay trip time constant in volt-seconds

UTP = ultimate trip point in volts rms, or voltage which will cause OVR to trip after being applied for an infinite time.

E_T = the transient rms voltage. Only the region $E_T > UTP$ is of interest.

T = time elapsed in seconds between the application of E_T and the energizing of the OVR relay.

Values of *UTP* and *K* will vary somewhat depending upon whether one, two, or three phases are high. If a constant value of E_T is considered, then the settings for the Electra are

$(E_T - UTP) = 10$ volts
$K = 4.5$ volt-seconds

with three phases high. These values represent the minimum that can be tolerated without risking nuisance tripping.

Current for the bias winding *F3-F4* is obtained from the voltage reference tube *V1* through current-limiting resistor *R15* (in actual practice the OVR bias winding is in series with the UVR bias winding and the bias current is obtained through *R14* and *R15*). The ultimate trip point is set by *R2* and the time constant by *R6*. The OVR relay is de-energized at normal voltage and energized by overvoltages.

By referring to Fig. 19, it will be seen that the overvoltage protection is accomplished in a manner similar to under-

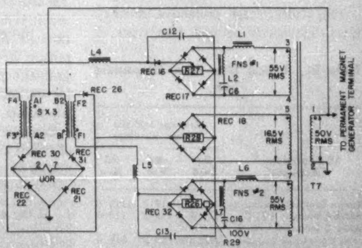

C12
L1
L4
REC 16
R27
FNS #1
L2
C6
55V RMS
REC 17
REC 26
F4
A1
B2
F2
SX3
F3
A2
B1
F1
REC 18
R28
165V RMS
50V RMS
TO PERMANENT MAGNET GENERATOR TERMINAL
REC 30
REC 31
2
7
UOR
L5
L6
REC 22
REC 21
FNS #2
R26
55V RMS
REC 32
L7
C16
C13
100V
R29
3
4
5
6
7
8
1
2
T7

Fig. 21. Under- and over-frequency relay diagram

voltage protection, in that a set of OVR contacts is used to energize the LOR, which in turn de-energizes the generator. Opening and then closing of the GCS will reset the protective circuitry and re-energize the generator. However, if the overvoltage fault remains, the overvoltage relay will again trip and de-energize the system.

UNDER- AND OVERFREQUENCY RELAY

Engine speed in flight is normally constant at 13,800 rpm. However, for ground idle the engine speed is reduced to 10,000 rpm. Therefore, in order to supply essentially constant frequency power for all conditions, a 2-speed gear-shifting device is incorporated into the airplane. Co-ordination with the gear-shifting operation requires the incorporation of the time delay *TDR3*.

Fig. 21 shows the frequency-sensing circuit used to control the under- and overfrequency protection circuitry. Essentially the circuit consists of a magnetic amplifier operated under- and overfrequency relay (UOR), two frequency-sensitive networks (FSN no. *1* and FSN no. *2*) and a reference network. The difference between the FSN's and the reference network outputs are fed through blocking rectifiers *R16* and *R26* to the two control windings of the magnetic amplifier. When no current flows in either control winding, which will be the case when the frequency is within the 360–425-cps band, the UOR relay will be energized. Outside this band, current will flow in one or the other winding and the relay will be de-energized. Frequency values at which current will start flowing in the control windings are determined by the settings of *R26* and *R27*.

Two FSN's are used because of the relatively wide band of frequencies over which the system must operate. If only one FSN were used, the curve of E_{out}/E_{in} versus frequency would have to be relatively flat to include both extremes. This would result effectively in a low milliamperes-per-cycle gradient for the magnetic amplifier control winding with a consequent large differential between relay pickup and dropout. With two networks, each one can be designed separately to give a high gradient in the relay trip region. Frequency response curves for both networks are shown in Fig. 22.

Consider now the operation of the UOR as frequency is varied from below 360 cps to above 425 cps. Starting below 360 cps, the output of FSN *1* and *2* will be greater than the output of the reference network. Hence current will flow through blocking rectifier *REC16* and winding *F3-F4*. UOR relay will be de-energized. No current will flow in *F1-F2* because of the position of *REC26*. Above 360 cps, the output of FSN *1* will be less than the output of the reference network and no current will flow in *F3-F4*. The output of the reference network will still be less than that of FSN *2* so no current will flow in *F1-F2*. The UOR relay will now be energized. Above 425 cps, the output of the reference network will be greater than that of FSN *2* and current will flow through blocking rectifier *REC26* and winding *F1-F2*, thus de-energizing the UOR relay.

The protective circuitry is shown in Fig. 19. It will be noted that a set of normally open UOR contacts and a set of normally closed auxiliary frequency relay (AFR) contacts are in series with *TDR3* coil. The *TDR3* relay is of the thermal timer type. Assuming that the UOR relay is initially de-energized, both the *TDR3* and AFR relays are de-energized. The AFR contacts in series with the output to the GCR coil will be open and the generator will be de-energized. Energizing the UOR will apply current from the PMG TR to the *TDR3* coil. After the 3 seconds' delay the *TDR3* contacts in series with the AFR coil will close. The AFR will be energized and lock itself in through its normally open contacts, at the same time, it will remove current from the *TDR3* coil allowing it to cool off and reset. The AFR contacts in series with the GCR coil will close and (assuming the GCR is closed) allow the field to be flashed.

FIELD FLASHING

Since the generating system is completely self-excited, a "flashing" source must be provided for initial generator buildup. The source of the flashing power is the PMG. The flashing circuitry is shown in Fig. 19.

Alternating voltage from the PMG is rectified by transformer *T6* and rectifiers *REC20* and *REC29* which produce a full-wave rectified signal of approximately 28 volts d-c at 6,000 rpm. Assuming that the generator is up to speed, the LOR is not energized, and the GCR is closed, current from the transformer-rectifier will flow through the GCR, normally closed LOR contacts and normally open AFR contacts to the GCR coil. This will lift the short circuits from the exciter SCPT potential windings and permit the exciter to function. Current for flashing will be fed to the generator field through current limiting resistor *R20* and the normally closed ACR contacts. As soon as the line voltage rises high enough to energize the UVR, the ACR will be energized and lock itself in thus removing the flashing power from the field. Blocking rectifier *REC27* is used to prevent the possibility of the field voltage feeding back to the GCR when the LOR is energized and the ACR is de-energized.

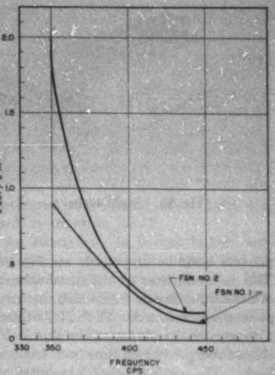


Fig. 22. Response of under- and overfrequency relay

Product Design of Electra Controls

It is not merely outstanding electrical performance that makes a unit acceptable for use in a modern aircraft, but reliability light weight, and maintainability as well. A major factor contributing to successful equipment embodying these characteristics is the simultaneous consideration of the problem of both installation and product design thereby providing the best over-all compromise. Major components within a unit usually determine its optimum shape and size. The shape, however, may or may not provide the optimum over-all result and, therefore, should be considered together. In addition, a willingness on the part of both equipment manufacturer and user to make modifications as required to acquire this goal is a must. The design of the equipment for this system is an example of this type of co-operation.

The static exciter weighing $24^1/_2$ pounds presented a major problem in face of the requirement that no shock mounts be used in meeting the specification vibration environment of 10 *g* up to 250 cps. The philosophy of design of a unit of this weight and size was to utilize the relatively stiff construction of the oblong-shaped SCPT as a backbone from which the few remaining com-

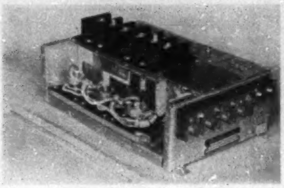


Fig. 23. Static exciter

ponents could be supported; see Fig. 23. This is in contrast to the usual scheme of building a chassis onto which the components are mounted. The SCPT utilizes C-type cores banded securely into a U-shaped member which runs from front to back. The terminal board is mounted on the top of this member and with the connections to the transformer make an integrated assembly. Because of its location in the main bus of the system, maximum protection from line faults within the exciter is provided. This is accomplished by using a rigid copper bus assembly flat against the U-member and sandwiched tightly between flat sections of silicone-glass laminate which insulate it from ground and against faults caused by foreign objects falling against it. Full-height barriers for all terminal boards are provided for the same reason.

The six silicon rectifiers are mounted onto flat radiators which in turn are mounted horizontally along the side of the transformers assembly adjacent and below the bus bars. Thus, with the perforated cover removed, all rectifiers and the bus assembly are readily accessible. The three linear reactors are mounted on the same member as the transformers, adjacent to one another across the front. The control and field terminal boards are mounted on the front. The main terminal boards on the exciter are made such that the cables to the transfer box can be connected from either right or left, thereby allowing identical units to be used for both sides.

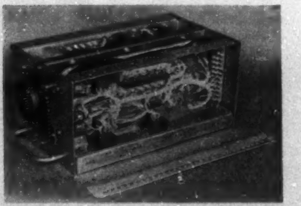


Fig. 24. Control unit

An example of the afore-mentioned modifications to meet the desired over-all objective was a change at the prototype stage from an exciter having the main terminals at the front and ATR-type mounts to the foregoing design. The ATR mounts with pins in the rear and clips in front were abandoned as being inadequate for holding down an unshock-mounted unit. They imposed an undue weight penalty as well in this application.

The control unit is much more conventional in design; see Fig. 24. This unit which resembles electronic rather than power equipment, also must survive in a 10 *g* up to 250-cycle environment without shock mounts.

The best technique for accomplishing this is the putting of complete functions into sturdy modules. This was considered but not used because of the economies of replacement; a $25 to $50 limit on unrepairable subassemblies being imposed by the user airlines.

Printed circuits were also considered and a prototype was built which used two 8-inch by 5-inch boards; however, this, too, was abandoned because the large number of bulky components such as relays, transformers, magnetic amplifiers, etc., made a dense package difficult. The problem of damage to the print circuit during component replacement was also a factor in the decision.

The construction used involved three full-length pan-type chassis, one on each side and one at the bottom. These combined with a U-shaped top cover and a flat bottom cover comprise rigid box members which withstand the vibration. Each chassis contains one or more complete functions such as an overvoltage relay assembly which expedites trouble-shooting or replacement. Solder terminal boards are provided on each panel for this purpose. Plugs, while of greater convenience, are not used because of their lesser reliability.

Components are mounted on the pan chassis in the conventional manner but with the wiring portion facing the outside where it is completely accessible when the covers are removed. Each chassis can be swung out at right angles to the panels main axis for access to both sides since the wiring harness feeds to a single terminal board on each end. Important test points are brought out to their same terminal board.

Over-all dimensions of the control unit are 6 by 7 by 11³/₄ inches and it weighs 13 pounds; see Fig. 25.

The small size of the regulator (6 by 6 by 5¹/₂ inches and 5¹/₂ pounds) allows use of a still more conventional pan-type chassis although it is made L-shaped to give greater mounting area and rigidity. The terminal board is the same as used on the front of the exciter.

Fig. 25. Voltage regulator

Appendix. Solution for Optimum X_d

$$X_d^2(X_d+a)=b$$

Since

$$\frac{b^2}{4}+\frac{a^3}{27}>0$$

(*a* is always +) then:

$$X_d=\pm 2\sqrt{\frac{a}{3}}\cot 2\phi$$

where

$$\tan\phi=(\tan\psi)^{1/3}$$

$$\cot\ 2\psi=\sqrt{\left(\frac{27}{4}\right)\left(\frac{b^2}{A^3}\right)}$$

For small values of ψ (i.e. small a),

$$\tan 2\psi\rightarrow 2\psi$$

and so

$$2\psi=\sqrt{\frac{4a^3}{27b^2}}\qquad \psi=\frac{1}{b}\sqrt{\left(\frac{3}{a}\right)^3}$$

Then by trigonometric identity:

$$\cot 2\phi=\frac{1-\tan^2\phi}{2\tan\phi}=\frac{1-(\tan\psi)^{2/3}}{2(\tan\psi)^{1/3}}\rightarrow$$

$$\frac{1-\psi^{2/3}}{2\psi^{1/3}}=\frac{1-\frac{1}{b^{2/3}}\left(\frac{a}{3}\right)}{2\left(\frac{1}{b}\right)^{1/3}\sqrt{\frac{a}{3}}}$$

So

$$X_d\rightarrow(b)^{1/3}\left(1-\frac{a}{3b^{2/3}}\right)$$

Reference

1. Some Applications of Magnetic Amplifiers in Aircraft Generator Protective Systems, D. L. Plette, J. W. Butler. *AIEE Transactions*, vol. 76, pt. II, 1957 (*Paper no. 57-481*).

Salient-Pole Permanent-Magnet Alternators for High-Speed Drive

ALLEN T. PUDER
NONMEMBER AIEE

FRITZ STRAUSS
MEMBER AIEE

MODERN prime movers, especially turbines, and new applications such as missile auxiliary-power units, have established requirements for reliable, efficient, high-speed machines for the generation of electric energy. The specific feature of the permanent-magnet alternator, i.e., elimination of the wound field and, hence, external excitation and slip rings, makes this type of machine superior to conventional machines for these applications. Reliability is enhanced by the elimination of arcing contacts which cause radio interference, brush operating difficulties at high altitudes, and explosion danger in confined, fuel-laden atmospheres.

This paper is concerned with a description and the design problems of some of the more recent units which have been built and tested in production. Rated outputs up to 10 kva have been realized at speeds to 24,000 rpm. Greater outputs and higher speeds have been achieved in experimental machines.

Nomenclature

B_g = air-gap flux density
B_{gm} = air-gap flux density reduced to magnet area
B_{sc} = air-gap flux density at short-circuited stator winding
B_{st} = flux density of magnet at point of stabilization
C_m = armature reaction factor[1]
C_p = ratio of average to maximum of the field form
C_1 = ratio of the fundamental to the actual maximum field form[1]
D = stator outside diameter
d_r = rotor diameter
E_t = alternator line voltage
E = alternator electromotive force (emf) per phase
E_{sc} = alternator emf at short circuit
F_{dm} = armature reaction, ampere-turns per pole
g_e = effective air gap
H_{dsc} = demagnetizing force of stator field (ampere-turns per inch magnet length per pole) at short circuit
H_{st} = demagnetizing force in the magnet (ampere-turns per inch magnet length) at the point of stabilization
h = width of stator core
I_{sc} = stator current per phase at short circuit
k = design constant
k_d = distribution windings factor
k_1, k_2 = proportionality factors
l = axial length of stator iron
m = number of phases
n_s = total number of effective stator conductors
P_g = air-gap permeance
P_{ge} = combined permeance of air gap and stator core
P_i = in-stator leakage permeance[2]
P_m = spatial permeance of permanent magnet[2]
P_o = out-stator permeance[2]
p = number of poles
r = resistance of stator winding per phase
x_l = stator leakage reactance per phase
z_{sc} = stator short-circuit impedance per phase
β = angle of stabilizing shearing line
γ = angle of working shearing line
δ = angle of "in-stator" shearing line
μ_e = effective permeability
μ_{st} = slope of stabilizing shearing line
μ_p = permanent permeability, slope of minor hysteresis loop[2]
σ = stacking factor of stator laminations
τ = ratio of reluctances of back iron and air gap

Mechanical Design

Basic design objectives for developing the permanent-magnet alternator for high-speed applications are

1. Maximum reliability under extreme storage and operating environmental conditions.
2. Minimum weight and space for a maximum output.

With these objectives in mind, alternators have been designed to couple directly to high-speed prime movers. This arrangement eliminates the requirement for any complex power transmitting components and minimizes the number of supporting bearings required.

These alternators are radial air-gap machines, of compact design and high efficiency. The stationary armatures of the machines are of conventional type, with single- or multiple-phase windings. The rotating fields utilize high-energy *Alnico V* permanent magnets. Fig. 1 shows a typical rotor construction, designed in a basic cage-type structure, which is die-cast with an aluminum alloy. The cage structure is built up of stainless-steel (nonmagnetic) plates, with rivets holding the laminated pole tips, *Alnico V* permanent magnets, and an inner magnet tie block in place around the shaft.[2] In a 2-pole rotor structure, the inner magnet tie block and through shaft are eliminated and flanged drive hubs are used, as shown in Fig. 2.

An interference fit is used in the cage assembly to support the magnet and to maintain a tight magnetic circuit. The magnet is used as a structural member in compression only, since *Alnico V*, with its exceptionally good magnetic properties, is a hard and brittle material with low tensile strength.

Die casting of the preheated rotor assembly is performed using a special casting die into which molten aluminum is forced from one end of the die to the other. High die-casting pressures ensure filling the mold and all internal voids, including the volume between the magnets and the damper-bar openings, with a dense material. Aluminum is used, since it is readily cast within the required temperature range, has good mechanical strength properties, and is a good electric conductor. Being an electric conductor, the aluminum serves as a closed turn, surrounding each magnet for the protection of the permanent-magnet field under transient short-circuit conditions. It also improves the waveform through the damper-bar winding and, in the case of an unbalanced load, reduces the unbalance of the phase voltages.

All critical parts used in the rotor structure are made from dies, in a punch-press operation, for economy in production quantities. The *Alnico V* magnets are cast in blocks, to provide maximum magnetic properties, and then ground to the finished dimensions. With all parts being identical for a given rotor diameter and number of poles, the stack height can be adjusted to meet a wide range of output ratings, using a minimum stock of parts and special dies.

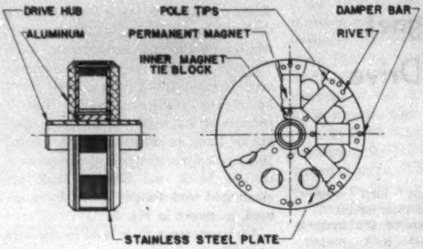
High-Frequency Alternator Design

The output frequency of permanent-magnet alternators, operating at speeds up to 24,000 rpm and constructed as described, has been limited in a small rotor by the impossiblity of magnetizing many closely spaced poles of opposite polarity. This limitation is eliminated by the unique rotor construction shown in Fig. 3. By alternating the permanent

---

Paper 57-489, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted January 15, 1957; made available for printing May 17, 1957.

ALLEN T. PUDER is with the Ruckstell Corporation, Azusa, Calif., and FRITZ STRAUSS is with the Christie Electric Corporation, Los Angeles, Calif.



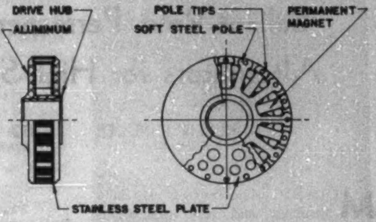


Fig. 3. Twenty-pole rotor construction with alternate permanent magnet and soft-steel poles

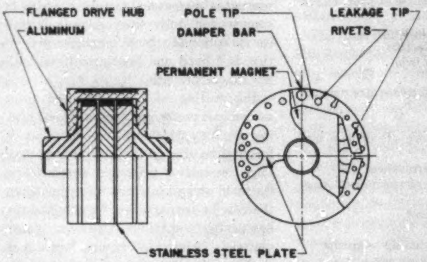


Fig. 1 (above, left). Eight-pole rotor, sectioned to show parts of typical cage structure with hollow shaft

Fig. 2 (left). Two-pole rotor, sectioned to show flanged drive hub and other parts of cage structure

magnet and soft-steel poles around the rotor circumference, so that the permanent magnets are of the same polarity, the entire rotor can be magnetized by a unidirectional radial field.[4] The soft-steel poles are saturated during the magnetizing impulse and immediately change to their proper polarity, becoming poles of a normal fully magnetized rotor. A typical small rotor of this construction, with its magnetizer, is shown in Fig. 4.

Stabilization

The permanent-magnet field of the alternator must be stabilized in order to avoid any voltage change subsequent to the application of a specified load condition. Stabilization methods commonly used, and used in the salient-pole alternator, are: 1. load stabilization, 2. short-circuit stabilization, and 3. air stabilization. The method or degree of stabilization used depends upon the requirements of a particular application. A brief description of each method follows.

LOAD STABILIZATION

Load stabilization is used in applications where weight is of utmost importance. Minimum weight is achieved by retaining the maximum generating flux. To accomplish this the rotor is assembled by means of a special magnetic keeper and then stabilized by applying the specified maximum load. Loads exceeding this stabilizing load will continue to decrease the operating voltage until a short-circuit condition is reached, beyond which no further demagnetization can take place.

SHORT-CIRCUIT STABILIZATION

Where a short circuit may accidentally be applied to the alternator, short-circuit stabilization is used. The rotor is assembled as described in the foregoing under "Load Stabilization." A sudden short circuit is then applied to the armature winding, while the alternator is operating at rated speed. Since a sudden short circuit causes the greatest possible load demagnetization, there will be no further decrease in the operating voltage, unless the alternator is disassembled.

DESIGN CALCULATIONS

The point of stabilization on the demagnetization curve of the permanent magnet, as well as the air-gap density, can be obtained by consideration of the following conditions:

1. The magnetomotive forces of rotor and stator must be in balance.
2. The sum of the voltages in the short-circuited stator winding must be zero.

As shown in the appendix, the air-gap flux density at no load, reduced to the magnet cross section, can be expressed by

$$B_{gm} = B_{st}\frac{P_g}{\mu_{st}}\;\frac{\mu_p+\mu_{st}}{P_g+P_l+\mu_p P_m} \tag{1}$$

where the slope μ_{st} of the stabilizing shearing line is

$$\mu_{st}=\frac{P_l+P_g}{P_m}\;\frac{P_l+kz_{sc}}{P_l+P_g+kz_{sc}} \tag{2}$$

and the constant k is

$$k=\frac{945C_p m}{C_m C_l k_d{}^2(n_s/100)^2(\text{rpm})/1{,}000} \tag{3}$$

Three-phase short-circuit stabilization of 2-pole machines frequently results in

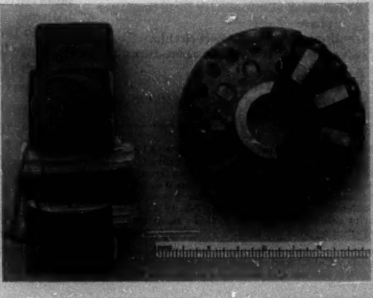


Fig. 4. Twenty-pole rotor with magnetizer which produces unidirectional radial magnetizing field. Alternator is 100 watts, 4,000 cycles per second, 24,000 rpm, designed to be mounted on a through shaft, rotor weight = 12 ounces

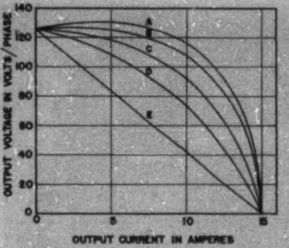


Fig. 5. Voltage-versus-current characteristic curves for a 1.5-kva 3-phase permanent-magnet alternator. Power factors:

A—0.95 leading
B—Unity
C—0.95 lagging
D—0.80 lagging
E—Zero lagging

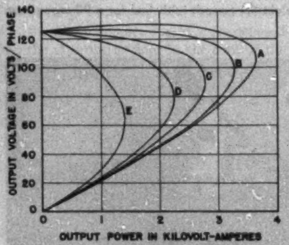


Fig. 6. Voltage-versus-power characteristic curves for a 1.5-kva 3-phase permanent-magnet alternator. Power factors:

A—0.95 leading
B—Unity
C—0.95 lagging
D—0.80 lagging
E—Zero lagging

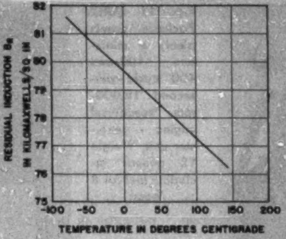


Fig. 7. Variation with temperature in residual induction of typical Alnico V permanent magnet

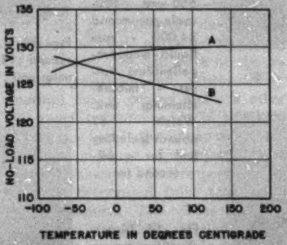


Fig. 8. Magnetic temperature compensation. No-load voltage over a temperature range of:

A—Overcompensated rotor
B—Noncompensated rotor

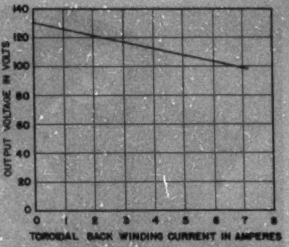


Fig. 9. Characteristic curve for typical toroidal back winding of no-load output voltage versus control direct current

a low stabilization point, because of the low in-stator permeance P_i. In such cases, it is recommended that the poles be provided with an additional leakage field, as shown in Fig. 2. The magnetic stabilization by short circuit renders the output voltage independent of demagnetizing effects of switching operations. This advantage, however, is only obtainable at the expense of increased weight.

For example, if the rotor has been magnetized in the stator, the initial working point would be located at the intersection of the line $\mu_{st}=(P_i+P_g)/P_m$, with the demagnetization curve. Application of load would bring the origin of the minor hysteresis loop to a lower point on the demagnetization curve. The exact location of this point depends upon the magnitude and the phase angle of the current. In the case of stabilization by short circuit, it is obvious from equation 2 that the stabilizing point on the demagnetizing curve moves downward until the aforementioned condition of balances between magnetomotive forces and between voltages is reached. Thus, the useful air-gap flux and, consequently, the voltage output will be reduced. To maintain the rated output characteristic of a machine not stabilized, an increase of the rotor volume will be necessary.

If a sudden short circuit is applied, the subtransient component of the short-circuit current will decrease the stabilizing point below the value expressed by equation 2. The degree of the change depends upon the leakage reactance of the damper winding, which should be kept at a minimum. A carefully designed aluminum structure, completely surrounding the magnets, can reduce the effect of the transient demagnetization to less than 3% of the output voltage of a machine stabilized by steady-state short-circuit current.

AIR STABILIZATION

Air stabilization is used in applications where frequent disassembly of the alternator is required without special tools. An air-stabilized rotor can be readily removed from the stator without a special keeper and yet maintain its output voltage after reassembly. A size and weight increase of 20 to 40% can be expected for an air-stabilized alternator having the same output characteristics as one which has been short-circuit stabilized.

Voltage Regulation and Control

Voltage control of small high-speed permanent-magnet alternators is a design problem because, in general, conventional types of control equipment have not been decreased in size and weight as rapidly as the alternators. Where feasible, the first design objective is to utilize the inherent regulation of the alternator. This approach eliminates all external controls. However, for lagging power-factor loads, or where close voltage control over a broad output range is required, the weight of the alternator with inherent regulation may become excessive, making external controls necessary. Figs. 5 and 6 show typical inherent regulation curves, based on a nominal rating which varies in accordance with the demands of a given application. Close inherent voltage con-

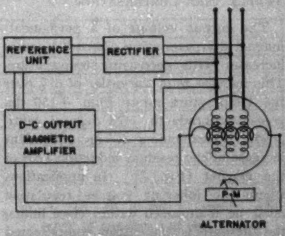


Fig. 10. Automatic control system using toroidal back winding for voltage control of permanent-magnet alternator

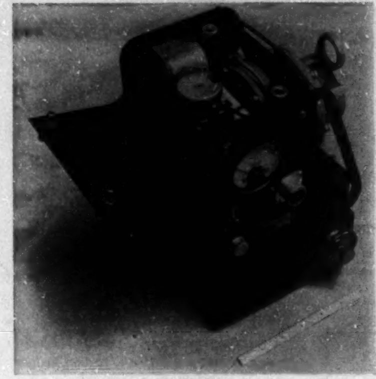


Fig. 11 (left). Pockette power plant, a miniature 100-watt 800-cycle-per-second 12,000-rpm semidiesel engine-generator unit. Weight 12 pounds including fuel for 8 hours

Fig. 13. Auxiliary power unit, 500-watt 400-cycle-per-second 12,000-rpm solid-propellant hot-gas-turbine-alternator unit. Weight 40 pounds with fuel for a 90-second run

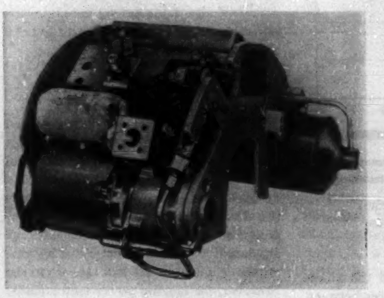


Fig. 12 (below, left). Accessory power supply, 4.0-horsepower 650-watt 400-cycle-per-second 24,000-rpm liquid-monopropellant-fuel hot-gas-turbine-alternator unit. Weight 47 pounds including fuel for a 60-second run

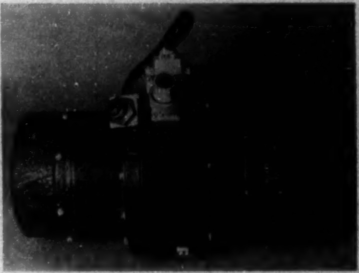


Fig. 14 (right). Auxiliary power unit, 1,500-watt 400-cycle-per-second 24,000-rpm compressed-gas-turbine-alternator unit. Weight 17 pounds (dry)

trol can be maintained over a maximum load range by: 1. establishing unity, or slightly leading, power-factor loads, or 2. increasing the power factor as step loads are applied. In the salient-pole alternator, temperature compensation and toroidal back winding are used for voltage control.

Temperature Compensation

The output voltage of a permanent-magnet alternator using *Alnico V* magnets increases with decreasing temperature. This increase is of the order of 6% over the temperature range from +150 degrees centigrade to −73 degrees centigrade, corresponding to an equivalent increase in the residual induction (B_r) of the magnet (Fig. 7). In applications where the inherent voltage regulation must be maintained within close limits, the alternator is temperature-compensated to eliminate this voltage change.

The temperature compensation is accomplished by substituting a temperature-sensitive (magnetic) metal for a portion of the stainless-steel (nonmagnetic) plates used in the rotor-cage structure. The function of this temperature sensitive ("thermomagnetic") alloy is to control, or shunt, a given amount of the usable generating flux. By providing a magnetic shunt made of a carefully selected thermomagnetic alloy[8], it is possible to compensate fully for the foregoing temperature effects. An alternator so compensated delivers a voltage which is essentially independent of temperature.

In practice, it is actually preferable to overcompensate in order to allow for the decrease in shunt resistance of the stator winding at low temperatures. Fig. 8 compares the no-load output voltage of an overcompensated rotor with that of a noncompensated rotor over a given temperature range.

Toroidal Back Winding

In applications where the inherent voltage of the alternator cannot be held within the required limits, a toroidal back winding is used for voltage control. This control winding is wound as a uniform-series toroidal winding around the armature core. The application of direct current to this winding causes a unidirectional field to be superposed on the alternating field in the armature core. As the superposed direct field is increased, the reluctance of the core increases, thus lowering the generated voltage. As shown in Fig. 9, the change in generated voltage is practically linear with the applied control current. Maximum control excitation is required under minimum or no-load conditions. Consequently, there is no increase of the armature heating, or decrease in the efficiency, under full load, as is the case in conventional alternators with field windings.

A typical automatic control system using the toroidal back winding is shown in Fig. 10. Static control systems of this type have been used to control the average value of the 3-phase output voltage within 1% over the complete load range.

Design of Toroidal Back Winding

It has been shown in the case of an air-stabilized machine,[3] that the density of the portion of the permanent-magnet flux crossing the air gap can be expressed by the equation

$$B_{gm}=B_{st}\frac{P_g}{P_0}\,\frac{P_0+\mu_pP_m}{P_g+P_l+\mu_pP_m} \tag{4}$$

where the air-gap permeance is

$$P_g=3.19\,\frac{\pi l}{2}\,\frac{C_pd_r}{g_ep} \tag{5}$$

The value g_e in equation 5 is a function of the air gap and the equivalent path of all soft-iron components of the magnetic circuit. Generally, the latter is so designed as to work essentially in the linear range of its magnetic characteristic.

If a direct magnetic field is superposed on the alternating field in the core of the stator, the permeance of the iron path will decrease. The ratio of the permeance of the air gap to the permeance of the stator core may be expressed as

$$\tau=P_g/P_b \tag{6}$$

where the permeance P_b of the stator core can be calculated by the equation

$$P_b=\frac{6.38}{\pi}\,\mu_e\,\frac{h\sigma lp}{D-h} \tag{7}$$

where μ_e represents the effective permeability of the stator core iron.

Substituting equations 5 and 7 in equation 6 yields

$$\tau=2.45\,\frac{C_pd_r(D-h)}{\mu_eg_eh\sigma p^2} \tag{8}$$

Replacing the air-gap permeance P_g in equation 4, by the equivalent permeance of air gap and stator core, $P_{ge}=P_g/(1+\tau)$, results in the reduced air-gap flux density at no load of

$$B_{gm}'=B_{st}\frac{P_g}{P_0}\,\frac{P_0+\mu_pP_m}{P_g+(1+\tau)(P_l+\mu_pP_m)} \tag{9}$$

Similar, for the case of short-circuit stabilization

$$B_{gm}=B_{st}\frac{P_g}{\mu_{st}}\,\frac{\mu_p+\mu_{st}}{P_g+(1+\tau)(P_l+\mu_pP_m)} \tag{9A}$$

The introduction of the effective permeability μ_e in equation 7 necessitates an explanation.

Upon entering the core, the rotor field subdivides in two opposite directions. Since the superposed direct field is unidirectional, the total field must be either the sum or the difference of the two, as the case may be. However, due to the nonlinear magnetic characteristic of the soft iron core, one half-wave of the resultant (polarized and synchronously rotating) flux is peaked and the other half-wave is flat-topped.[4] In addition, the rotating field in the air gap is far from sinusoidal in shape. Consequently, the required values of effective permeability must be determined from tests on a typical machine. The values thus obtained can be applied to design calculations of other machines if their core material is identical and the general geometrical configuration of their magnetic circuit is similar to that of the test machine.

The vectorial sum of the induced voltages in the conductors of the back winding is zero under ideal conditions. Irregularities in the winding and small unsymmetries in the magnetic circuit, however, may cause the build-up of a residual alternating voltage on the terminals of the direct back winding. This fact must be considered in connection with the design of the control circuit.

Applications

As a component in various types of power supplies, the salient-pole permanent-magnet alternator provides simple, rugged construction and the following significant advantages: 1. reliability with minimum weight, 2. high efficiency, 3. positive excitation, and 4. the absence of external excitation, which eliminates problems associated with moving contacts. Typical fields where these characteristics are required are in guided-missile auxiliary-power units, ground-support power supplies, emergency power supplies, and electric conversion units.

Fig. 11 shows a miniature, semidiesel engine-generator unit, using a small permanent-magnet alternator. When coupled directly to the engine crankshaft as a flywheel, a rugged and compact alternator rotor construction is essential.

A liquid-monopropellant-fuel hot-gas-turbine-alternator power supply is shown in Fig. 12. With the rotor directly coupled to the turbine, miniaturization is realized through high-speed operation. Operating on a monopropellant fuel, this unit provides electric and mechanical power completely independent of the missile's main power plant and fuel system. By temperature compensating the alternator rotor, the electric output voltage is held relatively constant over the extreme temperature conditions imposed by high-speed and high-altitude flight.

Another hot-gas turbine-alternator power unit, which uses a solid propellant, is shown in Fig. 13. This unit is a lightweight source of electric energy for air-borne applications. It consists essentially of a solid rocket propellant, an impulse turbine, an alternator, and an electromagnetic control system able to withstand long-term storage and insensitive to vibration, shock, and all normal air-borne environmental conditions.

A compressed-gas turbine-alternator

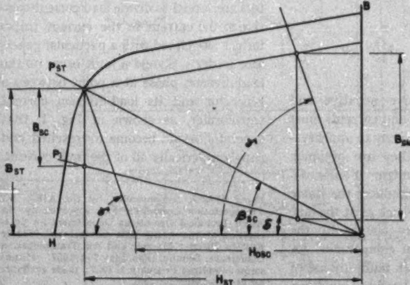


Fig. 15. Graphic representation of permanent-magnet circuit for stabilization by short circuit

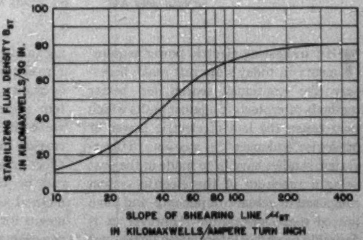


Fig. 16. Demagnetizing curve for Alnico V, represented as function of magnet flux density and slope of shearing line

power supply is shown in Fig. 14. Operational features of this power supply consist of close regulation of the speed and output voltage through the use of an electromagnetic speed-control system and the automatic toroidal back winding voltage-system.

Appendix

Air-Gap Flux Density for Short-Circuit Stabilized Machine

Fig. 15 represents the geometry of the permanent-magnet circuit.[3] It is assumed that the permanent magnets are stabilized on the demagnetization curve at the point B_{st}, H_{st}. The distance from this point to the intersection of its ordinate with the leakage line P_l is proportional to the air-gap flux density at short circuit B_{sc}.

Then, from Fig. 15

$$B_{sc}=B_{st}-H_{st}\tan\delta \qquad (10)$$

and the magnet flux density at the stabilization point is

$$B_{st}=(H_{st}-H_{dsc})\tan\gamma \qquad (11)$$

where H_{dsc} represents the demagnetizing stator ampere-turns per-inch magnet length in short circuit.

Substituting

$$\mu_{st}=\frac{B_{st}}{H_{st}} \qquad (12)$$

in equations 10 and 11 gives

$$B_{sc}=B_{st}(1-\tan\delta/\mu_{st}) \qquad (13)$$

and

$$B_{st}=H_{dsc}\mu_{st}\tan\gamma/(\tan\gamma-\mu_{st}) \qquad (14)$$

The hypothetical total magnetic flux in the air gap can be expressed by the equation[6]

$$\phi_r=B_g d_r \pi l=\frac{6{,}000E_t 10^8}{C_w n_s(\text{rpm})} \qquad (15)$$

where

$$C_w=\frac{E_t C_1 k_d}{E_p\sqrt{2}m} \qquad (16)$$

In the case of symmetrical short circuit ($E_p=E_{sc}$), the flux density B_{sc} corresponds to the remaining air-gap flux density, reduced to the magnet area,

$$B_{sc}=B_g\frac{C_p d_r \pi l}{b_p l_p p} \qquad (17)$$

Thus, from equations 15 through 17

$$B_{sc}=\frac{8{,}450C_p m 10^8}{b_p l_p p C_1 k_d n_s(\text{rpm})}E_{sc}=k_1 E_{sc} \qquad (18)$$

The induced emf E_{sc} must balance the impedance drop at short circuit

$$E_{sc}=\sqrt{r^2+X_{sc}^2}\,I_{sc}=z_{sc}I_{sc} \qquad (19)$$

The armature reaction can be expressed by[1]

$$F_{dm}=0.45n_s I_{sc} C_m k_d/p \qquad (20)$$

Substituting $F_{dm}=H_{dsc}h_p$, the stator current can be written

$$I_{sc}=h_p p/(0.45n_s C_m k_d)H_{dsc}=k_2 H_{dsc} \qquad (21)$$

Combining equations 18, 19, and 21 yields

$$B_{sc}=k_1 k_2 z_{sc} H_{dsc} \qquad (22)$$

Equating equations 13 and 22 and substituting equation 14 gives the slope of the stabilizing shearing line at short circuit

$$\mu_{st}=\frac{\tan\delta\tan\gamma+k_1k_2z_{sc}\tan\gamma}{\tan\gamma+k_1k_2z_{sc}} \qquad (23)$$

Introducing the following equation

$$k=k_1k_2P_m$$

$$=\frac{945C_p m}{C_m C_1 k_d^2(n_s/100)^2(\text{rpm})/1{,}000} \qquad (24)$$

and substituting

$$\tan\delta=P_l/P_m \qquad (25)$$

$$\tan\gamma=P_l/P_m+P_g/P_m \qquad (26)$$

gives

$$\mu_{st}=\frac{P_l+P_g}{P_m}\;\frac{P_l+kz_{sc}}{P_l+P_g+kz_{sc}} \qquad (27)$$

The stabilizing flux density B_{st} can be found directly for *Alnico V* from Fig. 16 after computing μ_{st} from equation 27.

Equation 4 has been derived on the basis of air stabilization at the outstator permeance, $P_o=P_m\tan\beta$. For stabilization at short circuit, equate $P_o=\mu_{st}P_m$ and obtain

$$B_{gm}=B_{st}\frac{P_g}{\mu_{st}}\;\frac{\mu_p+\mu_{st}}{P_g+P_l+\mu_p P_m} \qquad (28)$$

References

1. DESIGN CALCULATIONS FOR A-C GENERATORS, David Ginsberg. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1274–80.
2. SYNCHRONOUS MACHINES WITH ROTATING PERMANENT-MAGNET FIELDS, PART II. MAGNETIC AND ELECTRICAL DESIGN CONSIDERATIONS, Fritz Strauss. *Ibid.*, vol. 71, pt. III, Oct. 1952, pp. 887–93.
3. SYNCHRONOUS MACHINES WITH ROTATING PERMANENT-MAGNET FIELDS, PART I. CHARACTERISTICS AND MECHANICAL CONSTRUCTION, Maurice W. Brainard. *Ibid.*, Aug. 1952, pp. 670–76.
4. ROTOR CONSTRUCTION AND METHOD OF MAGNETIZING THE SAME, M. W. Brainard. *U. S. Patent no. 2,754,440*, July 1956.
5. TEMPERATURE-COMPENSATOR ALLOYS, Warren S. Eberly. *Machine Design*, Cleveland, Ohio, vol. 26, May 1954, pp. 152–56.
6. ALTERNATING CURRENT WAVE FORMS (book), Philip Kemp. Chapman & Hall, Ltd., London, England, 1952, pp. 279–81.

A Method of Fault Analysis for A-C Aircraft Electric Power Systems

L. J. RINDT
NONMEMBER AIEE

R. D. JESSEE
ASSOCIATE MEMBER AIEE

THE increasing electric power requirements in today's aircraft make necessary the continual search for better methods of protection against faults which may cause the loss of electric power or otherwise endanger the mission. Power-system studies have been made in the past using a mock-up of the electric system composed of actual components, and have proved quite successful. However, it is desirable to obtain information faster, and before the development of complete apparatus. Parameters such as line voltages and currents, positive-, negative-, and zero-sequence voltages and currents must be known in order effectively to apply system protection, and they are of much value in regulator design. Faults occurring where they produce the same effect on all regulators, such as at the bus, can be calculated very simply by the method of symmetrical components. A means for detecting such faults by use of negative-sequence voltage is pointed out in another paper.[1] However, in multigenerator systems, load-division circuits are employed in order that the load current may be properly divided between generators. Generator currents are usually sensed in a single phase through a current transformer; consequently, generators are forced to divide load current relative to the current in the current transformer associated with a particular generator feeder. Should a fault occur on the load-division phase at a point between a generator and its load-division current transformer, as shown in Fig. 1, that generator would become overexcited and supply practically all of the fault current.

Paper 57-479, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 1, 1957; made available for printing May 17, 1957.

L. J. RINDT and R. D. JESSEE are with the Westinghouse Electric Corporation, the former at East Pittsburgh, Pa., and the latter at Lima, Ohio.

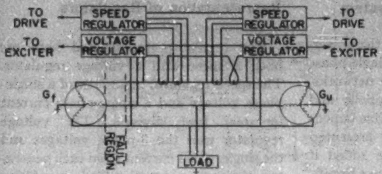


Fig. 1. Diagram of regulated system

Such a condition would be very difficult to calculate by ordinary methods.

The method of analyzing this type of system as presented here is to set up passive elements of an a-c calculating board to represent the elements of the power system and to use an electronic differential analyzer to represent the regulators of this system. Although there are transients which occur between the time the regulators are told to change until they have completed their change, the final steady-state conditions are the significant ones. At this steady-state point, the magnitudes and phase angles of the positive-, negative-, and zero-sequence components of the generator currents and voltages are needed. From these values, the generator power, voltages, and line currents can be computed.

In addition to evaluating this method, the effects of faults on a power system may be analyzed. With various system loadings, the effects of the following faults have been analyzed.

1. Single line-to-ground faults: The fault location is on the load-division phase at a point between the generator terminal and the load-division current transformer, as indicated in Fig. 1. The regulator receives its voltage signal from the load bus. Several different values of fault impedances are analyzed but in all cases the fault impedance is purely resistive.

2. Line-to-line faults: The fault location and impedances are the same as in the case of the single-line-to-ground fault. One of the faulted lines is the load-division phase.

3. Open-phase faults: The fault location is at a point between a generator and the load bus. This type of fault is applied to each phase with various amounts of load.

An analog-computer study has been undertaken to obtain data necessary to analyze these and evaluate the method.

Assumptions and Limitations

To set up the power system problem for analysis it is necessary to make certain assumptions and define limits imposed on certain parameters.

1. Saturation is neglected.

2. The exciter is assumed to be capable of supplying generator excitation demanded by the regulator sensing circuit from zero to ceiling.

3. Ceiling excitation voltage is assumed to be the same for all conditions.

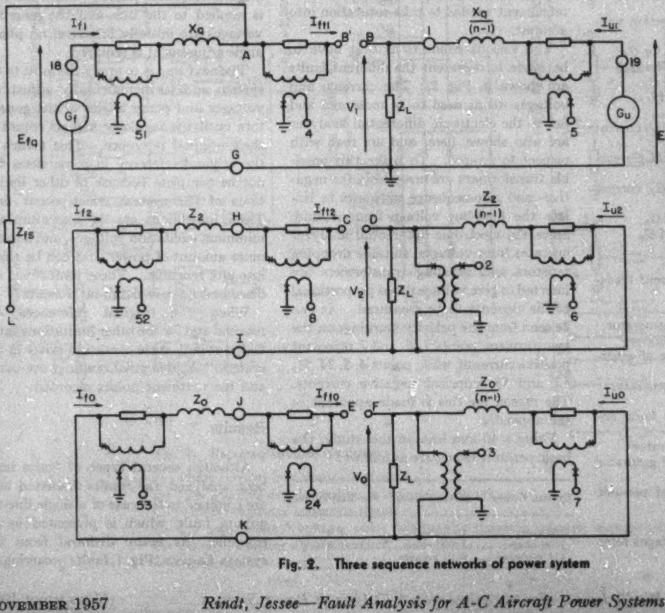
4. The generator is driven through an overrunning clutch which does not allow the prime mover to take power from the electric power system. It is, therefore, assumed that the maximum power taken from the system is 0.05 per-unit power, required to drive the clutch.

5. In any case where a generator is being motored and sufficient power cannot be supplied from the line to maintain synchronism, the generator is replaced by a constant impedance.

6. Frequency and temperature are assumed to be constant.

Nomenclature

G_u=unfaulted generator; represents $(n-1)$ generators
G_f=faulted generator
n=number of generators in system
E_{fq}=quadrature voltage of G_u
E_{uq}=quadrature voltage of G_f



TYPE OF FAULT	SHORT
SINGLE LINE TO GROUND	B' TO B C TO D E TO F G TO H I TO J K TO L
LINE TO LINE	B' TO B C TO D E TO F G TO I H TO L
OPEN PHASE	B' TO C TO E B TO D TO F

Fig. 2. Three sequence networks of power system

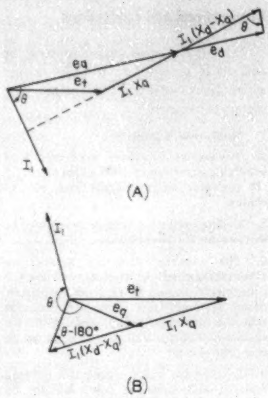


Fig. 3. Excitation phasor diagram

A—Maximum excitation
B—Minimum excitation

e_q = quadrature voltage of generator
e_d = excitation voltage of generator
e_t = terminal voltage of generator
V_1 = positive-sequence voltage at load bus
V_2 = negative-sequence voltage at load bus
V_0 = zero-sequence voltage at load bus
V_A, V_B, V_C = load-bus phase voltages
E_{f0} = reactive load-division correction voltage for G_f
E_{u0} = reactive load-division correction voltage for G_u
E_{fr} = regulated reference voltage for G_f
E_{ur} = regulated reference voltage for G_u
I_{f1} = positive-sequence current of G_f
I_{f2} = negative-sequence current of G_f
I_{f0} = zero-sequence current of G_f
I_{ft1} = positive-sequence current of G_f current transformer
I_{ft2} = negative-sequence current of G_f current transformer
I_{ft0} = zero-sequence current of G_f current transformer
I_{u1} = positive-sequence current of G_u
I_{u2} = negative-sequence current of G_u
I_{u0} = zero-sequence current of G_u
I_{fa}, I_{fb}, I_{fc} = phase currents of G_f
I_{fta} = phase A current of G_f current transformer
I_{ua}, I_{ub}, I_{uc} = phase currents of G_u
I_1 = positive-sequence current of generator
Z_q = quadrature reactance of generator
Z_2 = negative-sequence impedance of generator
Z_0 = zero-sequence impedance of generator
Z_L = load impedance
Z_{fs} = sequence representation of fault impedance
X_d = direct-axis reactance of generator
X_q = quadrature-axis reactance of generator
θ = angle between e_q and I_1
M = voltage-regulator constant of per-unit voltage/per-unit current

All sequence current and voltages refer to phase A quantities.

Representation of Power System

In this method of analysis, the power-system portion is set up on a single-phase basis using the three sequence networks, as shown in Fig. 2. For simplicity and because of their smallness, the line impedances are neglected in the case presented here. However, they may be added if their size indicates they will have a significant effect on the results. The generators G_f and G_u supply the voltages E_{fq} and E_{uq}, behind quadrature reactance. This particular representation was chosen because of two factors: 1. The generator can be represented by a single reactance and a single voltage source. 2. The voltage source used has the same phase angle as the excitation voltage, thus simplifying the calculations needed to determine the excitation voltage of the salient-pole machine.[2] A phasor diagram showing the relationship of the different voltage and voltage drops is shown in Fig. 3, and the method used to evaluate the excitation voltage is discussed under "Results."

As is often done in power-system analysis, the machine saturation is neglected. This will have no effect on the results where the system regulates within the range of the excitation voltage, but it will have some effect on the results where the upper limit of this range is reached. However, it is felt that this effect is not important enough to justify the additional refinement needed to take saturation into account.

The various connections that have to be made to represent the different faults are shown in Fig. 2. The currents and voltages which need to be measured and fed to the electronic differential analyzer are also shown here, and are read with respect to ground. To make this possible transformers are inserted in the negative- and zero-sequence networks to isolate the load-bus voltage from ground. Since the electronic differential analyzer operates from voltages, suitable dropping resistors and isolating transformers are inserted to give voltage drops proportional to the currents to be measured. As can be seen from the polarity markings on the transformers, points *5*, *6*, and *7* represent positive currents, while points *4*, *8*, *24*, *51*, *52*, and *53* represent negative currents. The reason for this is made apparent in the appendix.

Using a 40-kva base in this study, the basic per-unit values are as follows:

System Values	Board Values	Per-Unit Values
120 volts	10 volts	1.0
1.09 ohms	1,000 ohms	1.0
111 amperes	0.01 ampere	1.0

Representation of Regulators

The regulators represented here are both a speed and a voltage regulator. The speed regulator makes use of a single-phase voltage and a single-phase current for real load division. The voltage regulator uses the 3-phase voltages and the single-phase current from each generator through a reactive-load equalizing circuit. A description of these regulators appears in the appendix.

Procedure

In general, the procedure used in each of the studies is as follows: With no load or fault on the system, both generator voltages are brought in phase and set equal in magnitude to 1.0 per unit. This condition gives a bus voltage of 1.0 per unit with no current flowing between generators. Under this condition, voltage references E_{fr} and E_{ur}, Fig. 4(E), which the voltage regulators attempt to hold constant, are established. As a check, the speed regulator should indicate zero. These reference points are used througout all of the studies to indicate whether a regulated condition exists. The next step is to add the required load to the system. The voltages of the two generators are then readjusted until the two voltage-regulator signals have returned to their original reference points. Since the load is applied to the bus, and the generator voltages are initially in phase, no phase-angle adjustment is required.

The next step is to apply the fault to the system and to incrementally adjust the voltages and phase angles of the generators until the regulator signals return to their original reference. This return to the original reference in some cases will not be complete because of other limitations of the system which occur first. These limitations are the maximum and minimum excitation voltages, and a maximum amount of power that can be taken into any machine. These limitations are discussed more in detail in "Results."

When the original references are reached and/or the other limitations satisfied, a steady-state condition exists in the system. At this point readings are taken and the pertinent points recorded.

Results

Although several types of faults have been analyzed, the results discussed here are confined to the case of a single line-to-ground fault, which is presented as an example. As many deduced from the system diagram, Fig. 1, faults occurring on

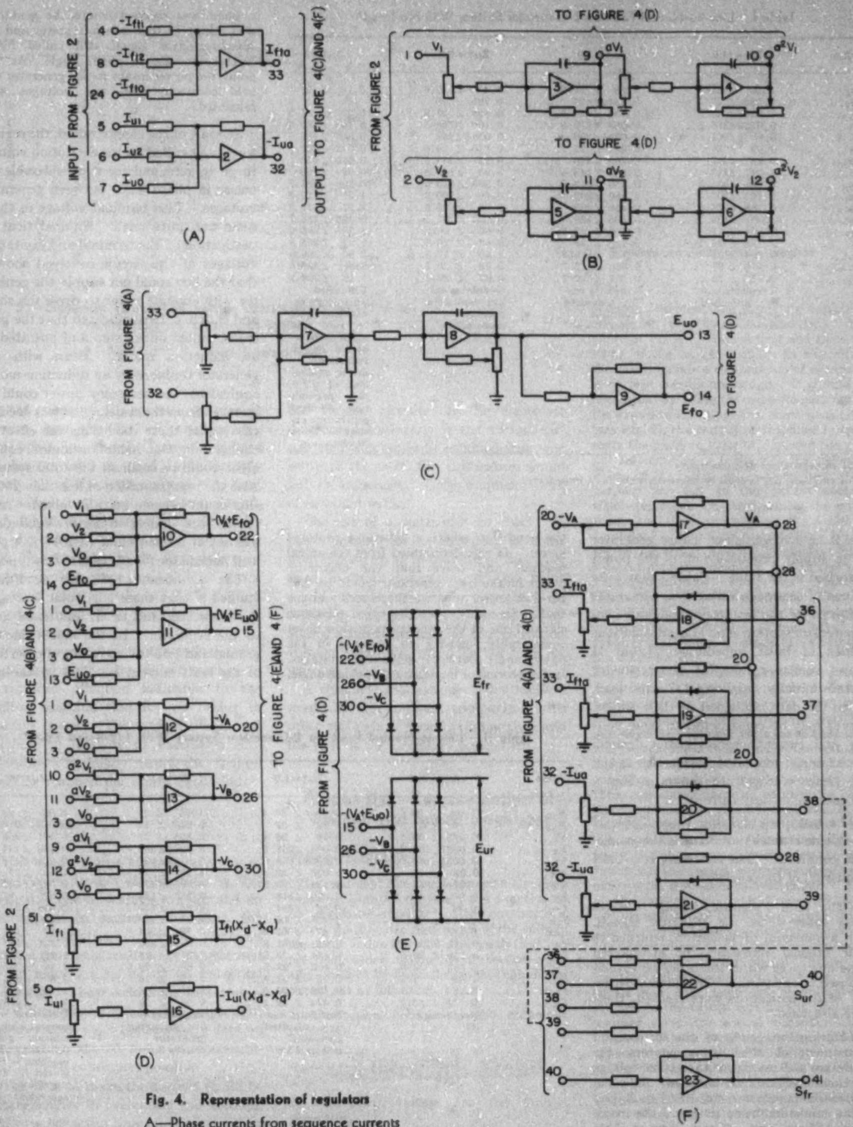

INPUT FROM FIGURE 2
4 -Ift1
8 -Ift2
24 -Ift0
Ift a 33
5 Iu1
6 Iu2
7 Iu0
-Iua 32
OUTPUT TO FIGURE 4(C) AND 4(F)
(A)
TO FIGURE 4(D)
FROM FIGURE 2
1 V1
9 aV1
10 a2V1
TO FIGURE 4(D)
2 V2
11 aV2
12 a2V2
(B)
FROM FIGURE 4(A)
33
32
Euo 13
14 Efo
TO FIGURE 4(D)
(C)
FROM FIGURE 4(B) AND 4(C)
1 V1
2 V2
3 V0
14 Efo
-(VA+Efo) 22
13 Euo
-(VA+Euo) 15
-VA 20
10 a2V1
11 aV2
-VB 26
9 aV1
12 a2V2
-VC 30
TO FIGURE 4(E) AND 4(F)
FROM FIGURE 2
51 If1
If1(Xd-Xq)
5 Iu1
-Iu1(Xd-Xq)
(D)
FROM FIGURE 4(D)
-(VA+Efo) 22
-VB 26
-VC 30
Efr
-(VA+Euo) 15
-VB 26
-VC 30
Eur
(E)
FROM FIGURE 4(A) AND 4(D)
20 -VA
VA 28
28
33 Ifta
36
20
33 Ifta
37
20
32 -Iua
38
28
32 -Iua
39
36
37
38
39
40 Sur
40
41 Sfr
(F)

Fig. 4. Representation of regulators

A—Phase currents from sequence currents
B—120- and 240-degree rotation of sequence voltages
C—Voltage-regulator correction signals
D—Corrected phase A voltage; phase voltages from sequence voltages; excitation component
E—Voltage-regulator signal to exciter
F—Speed-regulator signal

Table I. Line-to-Neutral Fault on 2-Generator System With No Load*

Value	$Z_{f0} =$ †3.0	Z_{j0}† = 1.0	Z_{f0}† = 0.6	Z_{f0}† = 0
E_{fq}	1.03 ∠ 42.3	1.23 ∠ 58.2	1.43 ∠ 67.6	2.1 ∠ 27
E_{uq}	0.985 ∠ 27	0.855 ∠ 27	0.705 ∠ 27	
V_1	0.995 ∠ 25.1	0.935 ∠ 25.8	0.855 ∠ 27.2	0.225 ∠ 1.0
V_2	0.0252 ∠273.7	0.068 ∠270.5	0.100 ∠268.5	0.190 ∠187.3
V_0	0.0035 ∠261.8	0.0096 ∠260	0.0142 ∠258	0.027 ∠176
V_A	0.99 ∠ 25	0.91 ∠ 22.2	0.815 ∠ 20.6	0.0225 ∠292.3
V_B	0.98 ∠267	0.905 ∠270	0.92 ∠273.5	0.345 ∠265
V_C	1.01 ∠147	0.995 ∠148.4	0.95 ∠149.6	0.38 ∠100
I_{fa}	0.06 ∠354.5	0.136 ∠318.8	0.236 ∠312.5	0.396 ∠316.7
I_{fb}	0.164 ∠207	0.456 ∠202.7	0.66 ∠200.2	1.27 ∠115.2
I_{fc}	0.168 ∠206.5	0.460 ∠203	0.66 ∠200.6	1.27 ∠115.8
I_{ua}	0.06 ∠174.5	0.136 ∠138.8	0.236 ∠132.5	0.396 ∠136.7
I_{ub}	0.164 ∠ 27	0.456 ∠ 22.7	0.66 ∠ 20.2	1.27 ∠295.2
I_{uc}	0.168 ∠ 26.5	0.460 ∠ 23	0.66 ∠ 20.6	1.27 ∠295.8
I_{f1}	0.3o4 ∠ 24	0.980 ∠ 16.5	1.46 ∠ 12.8	2.86 ∠299.3
I_{f2}	0.176 ∠ 26	0.466 ∠ 22.7	0.68 ∠ 20.7	1.3 ∠296.5
I_{f0}	0.172 ∠ 25.5	0.462 ∠ 22.7	0.67 ∠ 20.3	1.28 ∠296.2
	Regulates easily; unfaulted generator motors 2 kw	Regulates easily; unfaulted generator motors 2 kw	Regulates easily; unfaulted generator motors 2 kw	Unfaulted generator's excitation goes to zero and machine is replaced by an induction-motor equivalent of 0.4+j0.4 per-unit ohm. Faulted generator's excitation goes to maximum

---

* All values are per-unit quantities.

† For this case, Z_{f0} is equal to three times the fault impedance.

the load division phase cause generator G_f to supply practically all of the power absorbed in the fault. Table I shows the values of sequence voltages and currents obtained for the line-to-neutral fault on a 2-generator system at no load with various values of fault impedance. Table II shows similar results for a 2-generator system initially supplying 1/2 rated load. From the data contained in these tables, curves may be constructed to show voltage and current characteristics desired. For example, generator phase currents and bus phase voltages are shown in Fig. 5 as functions of fault current.

The last lines of Table I and Table II give some remarks concerning the particular condition. The following is a brief discussion of these remarks.

1. The term "regulates easily" denotes those cases where the regulator signals, with a minimum of oscillation, returned to their original references, or approached these references within the other limitations imposed on the system. In the cases where any of the limitations were reached, these were also noted.

2. During the course of the incremental adjustment of the two generators, the minimum and maximum excitation voltage on each machine was checked. This excitation voltage was $e_d = e_q + I_1(X_d - X_q)\sin\theta$, the minimum being zero and the maximum 4.85 per unit. When either of these limits was reached, the excitation voltage was held constant but the other variables were changed until the original reference or some other limitation was reached.

3. Another of the limitations reached was the condition where a generator motored power. It was determined from the actual machine and drive that the maximum power intake by a generator would be 0.05 per-unit power to drive the clutch and the shaft. In many cases the speed regulator signaled one of the generators to slow down considerably, causing an increase in the phase angle between generated voltages. In the process of increasing this phase angle, a point was reached where the generator was taking in 0.05 per-unit power and the speed-regulator signal still called for a further increase in phase angle. At this point, the power intake to the generator was held constant while the voltages were regulated.

In some of the cases studied, the regulator signals called for the excitation voltage to go to zero and for a considerable increase in phase angle between generated voltages. The terminal voltage in these cases was quite small. An analytical investigation of the terminal and excitation voltages of the circuit involved showed that the bus could not supply the generator with enough power to drive the shaft and clutch. This indicated that the generator pulled out of step and operated as an induction motor. Even with the generator replaced by an induction-motor equivalent, the necessary power could be taken from the system. It was decided that under these conditions the effect of changes in the induction-motor equivalent would be small on the total system. Also the representation of less than 200% slip would become quite involved. As a result, the induction-motor equivalent was set at a fixed value of 0.4+j0.4 per-unit impedance for all cases of this type.

The significance of the conditions studied is that these particular faults do not show the fault as an additional load to the regulator. In the single line-to-ground and line-to-line faults, the portion of the fault current supplied by the un-

Table II. Line-to-Neutral Fault on 2-Generator System With 1/2-Rated Load*

Value	Z_{f0}† = 3.0	Z_{f0}† = 1.0	Z_{f0}† = 0.6	Z_{f0}† = 0
E_{fq}	1.38 ∠ 54	1.7 ∠ 67.4	1.96 ∠ 72.5	2.03 ∠ 27
E_{uq}	1.18 ∠ 27	1.08 ∠ 27	0.945 ∠ 27	
V_1	0.975 ∠ 23.5	0.96 ∠ 26.4	0.895 ∠ 27	0.197 ∠359
V_2	0.0231 ∠269.6	0.066 ∠268.5	0.0995 ∠266.7	0.163 ∠184.6
V_0	0.0035 ∠260	0.0098 ∠259.6	0.0148 ∠257.7	0.0246 ∠175.3
V_A	0.98 ∠ 23	0.935 ∠ 22.5	0.85 ∠ 20.3	0.02 ∠291.4
V_B	0.97 ∠265	0.935 ∠269.5	0.865 ∠272.4	0.30 ∠262.1
V_C	1.0 ∠145	1.01 ∠148	0.985 ∠148.8	0.325 ∠ 97.6
I_{fa}	0.74 ∠359	0.84 ∠357	0.85 ∠350.2	0.536 ∠313.6
I_{fb}	0.174 ∠207	0.498 ∠205	0.74 ∠202.7	1.24 ∠117.3
I_{fc}	0.164 ∠205.3	0.474 ∠203.6	0.70 ∠201	1.17 ∠115.8
I_{ua}	0.316 ∠296.4	0.21 ∠281.8	0.095 ∠260	0.344 ∠133.5
I_{ub}	0.156 ∠ 22.5	0.454 ∠ 20.8	0.67 ∠ 18.5	1.15 ∠293
I_{uc}	0.163 ∠ 23.5	0.466 ∠ 22.5	0.69 ∠ 20.5	1.16 ∠295.1
I_{f1}	1.08 ∠ 7.3	1.75 ∠ 11.4	2.16 ∠ 10	2.8 ∠302
I_{f2}	0.162 ∠ 21.7	0.458 ∠ 20.5	0.67 ∠ 18.3	1.13 ∠293.5
I_{f0}	0.167 ∠ 23.1	0.474 ∠ 22.1	0.69 ∠ 20	1.16 ∠295.3
	Regulates easily	Regulates easily; unfaulted generator motors 2 kw	Regulates easily; unfaulted generator motors 2 kw	Unfaulted generator's excitation goes to zero and machine is replaced by induction-motor equivalent of 0.4+j0.4 per-unit ohm. Faulted generator's excitation goes to maximum

---

* All values are per-unit quantities.

† For this case, Z_{f0} is equal to three times the fault impedance.

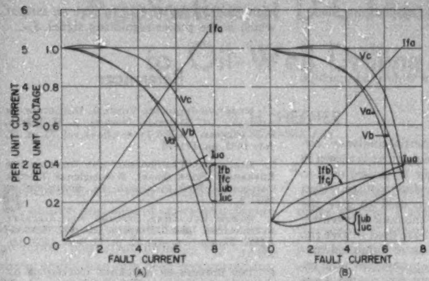


Fig. 5. Generator current and voltage characteristics with line-to-neutral fault on 2-generator systems

A—No load B—1/2-rated load

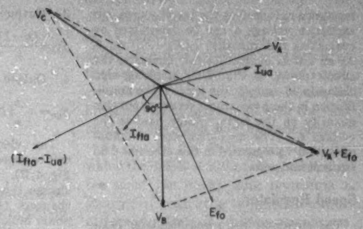


Fig. 6. Voltage-regulator phasor diagram

faulted machine G_u appears as current flowing into the faulted machine G_f. This effect, in general, causes the faulted machine to tend to speed up and supply nearly all of the fault current. As load was added to the bus, there was a tendency for the load current to be shared by the generators. In this case, and with a high-fault impedance, the unfaulted generator carried some load. With lower values of fault impedance, the unfaulted generator tended to slow down until it was motoring.

Fault current flowing through the current transformer of the faulted machine in a reversed direction indicated to the faulted machine that it was taking in reactive power. This caused the faulted machine to increase its excitation considerably. In many cases the faulted machine reached maximum excitation.

Conclusions

This study shows that an analysis of this type of system can be done by this method. The limitations are few and do not appear to seriously handicap the study. Limitations do occur in the cases where a generator needs to be represented by an induction motor or an induction generator. These conditions are extreme, but if their effects are deemed important, a proper analysis of the necessary data on the machine should make a representation of these conditions possible.

The time to make the studies would be another factor in evaluating this method. The original setup time on the computers was 2 days. The time for a run averaged about 90 minutes. However, both of these times could be shortened under the proper conditions. A considerable part of the setup time was used in calibrating the various circuits on the electronic differential analyzer. If the various circuit elements could be retained from one setup to the next, this calibration would not be necessary. Only a quick check of it would suffice.

The use of a wattmeter to check the power intake on the motoring machine would speed up all of the runs of this type, since a visual indication of this point would save much time consumed in adjusting the generators. A method of giving a direct reading of the excitation voltage would also speed up the time for various runs.

A complete evaluation of this method would involve a direct comparison with other methods of such factors as time and the cost of the necessary equipment.

Appendix. Representation of Voltage and Speed Regulator

Voltage Regulator

To represent the regulator with the electronic differential analyzer, the function of the regulator must be spelled out analytically. In this case, the circuit of the voltage regulator indicates that the analytical expressions which define the correction voltages supplied by the reactive-load equalizing circuit are as follows:

1. Correction voltage for the unfaulted generator

$$E_{u0} = -j\frac{M}{n}\left[I_{fta} - \frac{I_{ua}}{(n-1)}\right]$$

2. Correction voltage for the faulted generator.

$$E_{f0} = \frac{j(n-1)M}{n}\left[I_{fta} - \frac{I_{ua}}{(n-1)}\right]$$

$$E_{f0} = (1-n)E_{u0}$$

The electronic-differential-analyzer circuits to represent the foregoing equations are shown in Figs. 4(A) and (C). The value of E_{u0} is developed first and then the value of E_{f0} as $(1-n)E_{u0}$. In Fig. 4(A), the three sequence components of the phase A current signals are added. It should be noted that by using the opposite polarity on the transformer supplying these quantities (see Fig. 2), the output of amplifier *1* represents the positive value of I_{fta}, and the output of amplifier *2* represents the negative value of I_{ua}. Fig. 4(C) shows how these outputs are fed to the potentiometers, which represent the constants in the equations, and then summed as the input to amplifier *7*. The input to amplifier *7* is thus $[M/n][I_{fta} - I_{ua}/(n-1)]$. Amplifiers *7* and *8* each rotate this quantity -45 degrees for a total rotation of -90 degrees. This represents the $(-j)$ in the equation. The output of amplifier *8* represents the correction voltage E_{u0}. Amplifier *9* multiplies the value of E_{u0} by the constant $(1-n)$ to give the value for E_{f0}.

To make use of these correction voltages E_{u0} and E_{f0}, it is necessary to have the phase voltages of the load bus. Fig. 4(B) indicates the method of rotating the positive- and negative-sequence voltages so that they may be combined to produce phase voltages. Amplifiers *12*, *13*, and *14* in Fig. 4(D) are summing amplifiers which combine the three sequence quantities to give the three phase quantities. Amplifiers *10* and *11*, in this same figure, combine the phase A voltage with each of the correction signals. Fig. 6 is a phasor diagram showing the relationship of the various currents and voltages used in the voltage-regulator circuit. This diagram represents a typical case where $n=2$ and an unbalance in reactive power exists. The dashed lines on this diagram represent the distorted set of voltages that determine the regulator sensing signal during the reactive unbalanced condition. Fig. 4(E) shows how this same set of voltages is applied to the bridge rectifiers in the same manner as in the actual regulator. The outputs of these rectifiers are the regulator signals to the exciter, E_{fr} and E_{ur}.

The voltage regulator in this particular case has two limitations, maximum and minimum excitation voltage. This excitation voltage can be described by

$$e_d = e_q + I_1(X_d - X_q)\sin\theta$$

The phasor diagram representing this equation is shown in Fig. 3. Angle θ is the angle between the voltage e_q and the current I_1. Since the voltage behind quadrature reactance e_q is used as the board source, the excitation voltage e_d can be determined by evaluating the quantity, $I_1(X_d-X_q)\sin\theta$. As an aid in evaluating this quantity, the circuit shown in Fig. 4(D) involving sign changers *15* and *16* was used. The amplification in these two stages serves to put this quantity on the same voltage base as the voltage e_q. The maximum and minimum excitation voltage used in this case was 4.85 per unit and zero.

Speed Regulator

The circuit of the speed regulator indicates that the formulas for the power indication and the regulator output signal are:

1. Power indication for the faulted machine

$$P_{fr}=|0.324I_{fta}+V_A|-|0.324\,I_{fta}-V_A|$$

2. Power indication for the unfaulted machine

$$P_{ur}=\left|\frac{0.324\,I_{ua}}{(n-1)}+V_A\right|-\left|\frac{0.324\,I_{ua}}{(n-1)}-V_A\right|$$

3. Output signal to the faulted machine

$$S_{fr}=\frac{(n-1)}{n}(P_{fr}-P_{ur})$$

4. Output signal to the unfaulted machine

$$S_{ur}=-\frac{1}{n}(P_{fr}-P_{ur})$$

The electronic-differential-analyzer circuit to represent these equations is shown in Fig. 4(F). Sign changer *17* is used only to change the sign of the bus voltage. Sign changers *18* and *19* are used to give the power indication P_{fr} of the faulted machine. Sign changers *20* and *21* are used to give the power indication P_{ur} of the unfaulted machine. The input potentiometers and resistors of these latter four sign changers are used to set the coefficient of the current and to put it on the same voltage base as the bus voltage. The diodes in the feedback circuit introduce the absolute magnitudes. Amplifier *22* sums up these power indications, amplifies them to represent the constant $(-1/n)$, and gives an output which is the power-regulating signal S_{ur}. Amplifier *23* multiplies S_{ur} by $(1-n)$ to give an output which is the power-regulating signal S_{fr}.

References

1. Negative-Sequence Voltage Detection on Aircraft A-C Electric Systems, R. D. Jessee, P. F. Boggess. *AIEE Transactions*, vol. 75, pt. II, July 1956, pp. 171–75.

2. Electrical Transmission and Distribution Reference Book (book), Westinghouse Electric Corporation, East Pittsburgh, Pa., fourth edition, 1950, p. 175.

3. Power Equalizer Systems for Aircraft Alternators, John A. Granath, Albert K. Hawkes. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 209–17.

4. The Effects of Abnormal Conditions on Aircraft, Parallel A-C Power Systems, S. C. Caldwell, A. J. Wood. *Ibid.* (Jan. 1954 section), pp. 379–86.

5. Analog Computer Methods Applied to Steady-State A-C Problems, H. A. Kahle, H. M. McConnell. *Ibid.*, vol. 75, pt. II, Nov. 1956, pp. 279–82.

6. Symmetrical Components (book), C. F. Wagner, R. D. Evans. McGraw-Hill Book Company, Inc., New York, N. Y., 1933.

◆

LATE DISCUSSION

The following discussion was received too late to be included with the paper.

Protecting A-C Motors with Low-Voltage Air Circuit-Breaker Series Trip

Discussion of paper 56-652 by F. P. Brightman, R. R. McGee, and P. J. Reifschneider, published in Applications and Industry, July, 1957, pages 114–19.

E. P. Ries (Shell Chemical Corporation, Martinez, Calif.): My comments are best presented by first briefly summarizing the electrical systems in our newer petro-chemical plants and then telling how the information presented in the paper affects us. Plants considered are characterized by continuous process operations. Most equipment drivers range from 3 through 75 hp (horsepower) with a few at 100 hp and larger. Continuity of service to most is critical. Single- or double-ended selectively co-ordinated unit substations supply all 440-volt loads. Motors of 3 through 75 hp are controlled by conventional general-purpose control centers supplied through a feeder circuit breaker. Motors of 100 through 150 hp are controlled by individual large air circuit breakers.

This is in contrast to the older systems where unit substations were equipped with manually operated instantaneously tripped feeder circuit breakers. Motors of 3 through 75 hp were controlled by explosion-proof switchracks, each protected by a molded-case main circuit breaker. Motors of 100 hp and larger were controlled either by means of high-voltage contactors or power circuit breakers.

This means that we now expect the same selective tripping or co-ordination with the low-voltage system that we previously expected only from our higher voltage system.

Associated with this concept are the following problems which we feel deserve added attention:

1. Motor heating curves to permit proper application of trips are seldom available or obtainable.

2. Width of the trip tolerance band in some cases makes proper application either difficult or impossible.

3. It is necessary to obtain special factory calibration when close co-ordination is required.

4. Test equipment for an initial field calibration check is not available at an economical price.

5. For the same reason it is not possible to check the calibration at routine intervals as is done with a relay.

6. It would not always be possible to reset the trip properly in the field even if the necessary test equipment were available.

I should like to suggest that effort be directed toward the following:

1. Narrower trip tolerance bands.

2. Field test exuipment that can be justified by industrial users or redesign of trip units for use with doughnut-type current transformers so that normally available equipment can be used.

3. Provisions for complete field adjustment of trip units.

4. Publication of heating curves by motor manufacturers.

Application of D-C Machines to Oil-Well Drilling

B. H. HEFNER
NONMEMBER AIEE

BECAUSE OF their flexibility, portability, and ease of operation, diesel-electric units are particularly well suited as power plants for drillings rigs. The diesel engine has long been recognized for its efficiency, reliability, and economy of operation. In many cases, the diesel engine has all but replaced other types of prime movers.

An electric-transmission system is an infinitely variable speed changer that matches engine output to load requirements. It eliminates the need for mechanical-torque multipliers and multiple-chain drives. Instead, it utilizes flexible cables that can be connected and disconnected quickly and easily without requiring bolting and taping. The need for alignment of basic power units with draw works, rotary tables, or mud pumps is thus eliminated and set-up time is greatly reduced. Multiple loads can be handled by one engine without gear-reduction complexity. The flexible connections permit the engines to be located on a tender while the rig and its equipment are on a platform. One engine can also be mounted on the platform to provide it with independent power should it be necessary to move the tender because of weather conditions.

Diesel-electric rigs are light and inherently compact. The latter is particularly important to offshore operations where space is at a premium. On land, the need for heavy foundations is avoided and set-up time is cut to a minimum. Driller's controls for the electric-transmission system are as simple as, if not simpler than, controls for conventional transmissions.

Some versions of of diesel-electric units utilized conventional switchgear and off-the-shelf industrial motors. The equipment was large and controls were complex and costly. Rig operators had to turn to specialists to take care of even minor difficulties.

With the use of the General Motors *567*-series 2-cycle *V*-type 8-cylinder 875-hp (horsepower) diesel engine as the prime mover, the design and building of an electric-transmission system that would satisfy as closely as possible the following criteria was initiated:

1. Over-all system as simple as possible.
2. Maximum interchangeability of components.
3. Maximum operational flexibility.
4. Built-in protection for all components.
5. Ease of operation and maintenance.

Drill-Rig Load Requirements

Since any transmission, mechanical, electric, or otherwise, essentially functions to match engine output to load requirement, it is obvious that the load requirements of the drilling rig must be analyzed and compared with engine characteristics.

The *567* diesel engine at full load operates at constant speed and constant torque. Drilling-rig loads are varied and arise from the following:

1. Draw works.
2. Rotary table.
3. Mud pumps.
4. Miscellaneous loads including air compressors, motor-generator sets for a-c supply, small slush pumps, coring reels, etc.

The draw works is basically used as a hoist mechanism, although it is also used to power the breaking-out and making-up catheads and sometimes to drive the rotary table. Hoisting the drilling stem is essentially a tractive effort. It is necessary to start and to accelerate a heavy load when the stem is long and deep. It is necessary to operate at higher speeds and low torque when the load is light, i.e., when a single stand is being handled.

This traction loading can be seen clearly in Fig. 1 which compares the curve for draw-works hook pull versus hook speed, with the curve for locomotive tractive effort with locomotive speed. The two curves have almost identical shapes.

As currently built, draw works are equipped with either a friction clutch or a jaw clutch. To engage a jaw clutch, the transmission comes to a full stop, while a friction clutch engages "on the fly." In designing a transmission for a friction-clutch-equipped draw works, it is particularly important to provide good deceleration control so that pick-up speed is not excessive. A case in point would be breaking out or making up with the catheads—a low-speed light-load operation.

As long as the drilling conditions for which the driller sets the rotary table stay the same, the rotary table represents a steady load with little variation in torque or speed. But if drilling conditions change, for instance, if gumbo or heavy strata is encountered and the bit begins to ball, torque will increase and speed decrease. If an "easy" strata is encountered, speed will increase and torque will drop.

The characteristics of the load represented by mud pumps are similar to the rotary-table characteristics. As long as the conditions for which the mud pumps have been set do not change, pumping speed does not vary. But should a cavity, plugged hole, or other situation arise, a change in pumping speed is necessary.

The remainder of the load from auxiliary equipment and services represents a small percentage of the total power requirements of the rig and is essentially a steady-state load.

Based on a consideration of these load patterns, it would seem logical to proceed on the assumption that, among other things, the electric transmission for a drilling rig should provide extremely high torque at low speeds and, conversely low torque at high speeds. In general, the transmission of power should be responsive to changing-load conditions. These objectives were strived for in the design of the electric transmission for drilling units.

Generator Selection

To supply power adequately to the load motors on a drilling rig, generator voltage should decrease as the current increases. In other words, a differential effect is required. The classical means of accomplishing this is to provide a differential field in the generator. In the original design studies the use of a generator of this type was considered. Its construction and characteristics are quite familiar from work in the locomotive field. However, upon further analysis, it was decided not to use it for the following reasons:

1. Multiplicity of load functions cannot be met without serious control complications.
2. Differential-field windings occupy space in the generator and require bigger forward-excitation windings. This results in the

---

Paper 57-978, recommended by the AIEE Petroleum Industry Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference of the Petroleum Industry, Philadelphia, Pa., September 9–11, 1957. Manuscript submitted January 14, 1957; made available for printing August 7, 1957.

B. H. Hefner is with General Motors Corporation, La Grange, Ill.

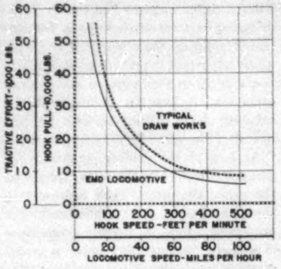


Fig. 1. Comparison curves of locomotive tractive effort versus draw-works hook pull

following: a heavier machine, and a larger machine.

It was decided, instead, to use a smaller generator with a separately excited field. The same machine, except for a change in the field winding, is used as a motor to power the draw works, mud pumps, and rotary table.

Designated as model *D39G*, the generator designed and built for drilling-rig application is a 4-pole machine rated at 500 kw. The armature has a progressive simplex lap winding and employs hot-pressed coils with a class *H* system insulation of glass and mica tape, glass and mica channels, and silicone varnish. Insulation on coil diamonds include varnished and melamine asbestos, fiber glass, and Teflon. Radial brush arrangement allows rotation of the armature in either direction. Main fields and interpoles have class *H* silicone-rubber insulation bonded to the copper which provides a positive moisture seal. Sealed-grease-lubricated cylindrical roller bearings at drive end and commutator end of the armature shaft provide extended maintenance-free service life.

The *D39G* generator is a force-ventilated machine, receiving its air from an engine-driven blower.

Excitation-Control System

In the design of the control circuit and apparatus, simplicity, rugged construction, and compactness were the objectives in the choice of components. An a-c power source was chosen for the control of generator excitation because it permits the use of minimum sized equipment requiring very low voltage and current. The excitation-control circuit is shown in Fig. 2 and is described in the following.

The first requirement for the control system is a device for the driller's hand that will send a command to the power plant to supply power for a given load. A number of devices can be used including a rheostat or a potentiometer. A Variac (a variable-voltage transformer) was used, which is a rugged dependable low-watt-loss unit that keeps at a minimum the heat-dissipation level at the driller's station.

The 110-volt a-c power supply, which is common to almost every rig in the field, is taken through the Variac. Controlled excitation flows from this point through a multiple-conductor cable to a motor-control cabinet, and from there to the generator-control cabinet on the engine skids by means of appropriate interlocks. The 110-volt a-c power goes through an isolating transformer which steps the voltage down by a 4-to-1 ratio. The current is then rectified by a bridge rectifier, a full-wave rectifier of the dry-plate type which is conservatively rated so that its life will be almost indefinite. The rectified current is then supplied to the field of the engine-driven exciter generator, which amplifies the excitation current to the main-generator field. The main-generator field requires up to 100 amperes.

Up to this point this is, in effect, a

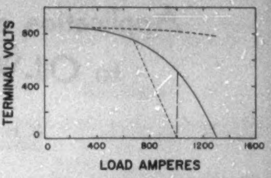


Fig. 3. Net results of the control system

shunt generator with the voltage dropping slightly with increased current demand. But to serve properly the constant-horsepower demand of the load motors, the control circuit must supply a differential effect so that voltage will drop as current demand increases. This is accomplished by means of a feedback circuit.

Feedback Circuit in the Control System

A feedback current proportional to the armature current of the main generator is supplied by a transductor which is inserted in the power circuit between the generator and motor. The device is essentially a type of d-c transformer. The d-c load cable from the generator to the motor passes through steel cores with coil windings around them. These windings are supplied by a small alternating current from the 110-volt supply, and a rectifier is added so that the feedback current is rectified.

Little current flows through the transductor until the d-c load cable begins to carry current. Then, as the load current increases, the transductor core saturates, changing its impedance and allowing an increase in the transductor output current. This current is in a fixed ratio to the armature current—approximately 1 to 400 amperes, and is unaffected by any other regulating action.

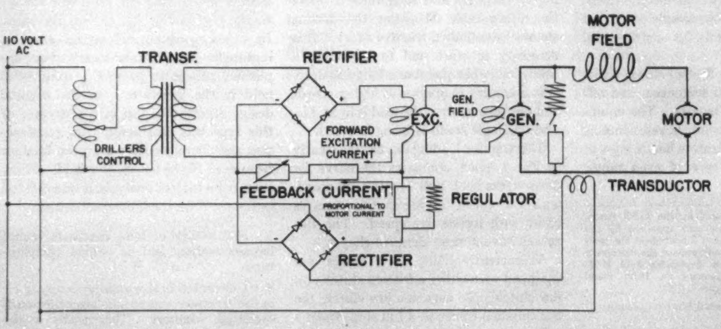


Fig. 2. Excitation-control circuit

The direct current from the transductor flows to a preset regulating resistor and a voltage drop is produced which subtracts from the voltage on the exciter generator field. The result is a reduction in exciter-field current and thus less voltage generated by the main generators.

Given a constant maximum voltage on the excitation circuit, the excitation current is determined by that voltage and by the resistance in the circuit. The feedback current is essentially independent of the resistance within its circuit and, as mentioned earlier, is proportional to the high-voltage d-c flow.

Thus, by using three simple low-cost static-control components, namely transductor, rectifier, and resistor, together with a relatively small generator, it is possible to obtain the characteristics of a much larger generator with a differential-field winding.

Regulation for Thermal and Torque Limits

With the addition of a simple field regulator two other important objectives are attained with the excitation-control system. These objectives include provision for torque and thermal limit. The torque limit is utilized for intermittent generator output to prevent damage to the equipment being driven. The thermal limit is utilized for continuous generator output to prevent overheating of the motors.

The field regulator has two control coils, a current coil and a voltage coil. The torque limit is provided by the current coil. The thermal limit is provided by the combined action of the current and voltage coils. Both act so that when energized, the operator on the resistance element moves to cut in resistance. When de-energized, the spring-loaded operator returns to its minimum resistance. The current coil is in series with the transductor output and becomes operative when the transductor output tells it that the armature (load) current is approaching 1,000 amperes. The coil then cuts resistance into the circuit and the transductor current passing through it creates a voltage which opposes the voltage of the excitation current.

The voltage coil, in series with a calibrating resistor, is wired across the generator output so that as the generator voltage rises the voltage coil becomes stronger, causing resistance to be cut into the exciter-field circuit. A recalibration relay is arranged so that when an individual generator is used to power an intermittent load such as the draw works, the relay cuts out the action of the voltage coil. This causes a change in the generator characteristic to match the operating requirements of the draw works.

The net results of the control system can be seen in Fig. 3 in which terminal volts have been plotted against load amperes to show the generator characteristic curves. The horizontal dashed line (at the top) is the generator characteristic resulting from a given excitation without control. The solid line shows the differential effect of the feedback current. It is a close approximation of the constant-power curve desired for drill-rig generator operation. The vertical dashed line (long dashes) represents the control effect of the field-regulator current coil on intermittent generator output to limit torque.

The diagonal dashed line represents the continuous generator output obtained as a result of the control effect of the current and voltage coils of the field regulator. As can be seen, the volt-ampere relationship is a straight line whose slope can be varied by varying the relationship between the current and voltage coils of the field regulator.

It can be seen that the generator-and-control system offers unusual flexibility in performance. Generator characteristics can be adjusted by relatively simple circuit changes to meet the sometimes widely varying needs of mechanical equipment used on drilling rigs.

Motor Selection and Control Design

The design approach used up to this point has not necessarily presupposed the use of a particular class of load motor. Rather, the choice of motor is dictated by the operational requirements of the drilling rig.

There are only two basic types of d-c motors, the shunt motor and the series motor. A variation of the shunt motor is the separately excited motor. A shunt motor with armature and fields connected in parallel with a common supply is immediately ruled out in view of the many high-starting-torque loads, such as in the case of the draw works. A shunt motor has a constant speed and a varying torque. It has very poor starting-torque characteristics. This type of motor requires added control such as accelerating resistors and contactors in the armature circuit or a separate excitation supply to provide the desirable starting characteristics.

The net result is a choice between a series motor and a separately excited motor. The series motor is widely accepted as the standard class of motor for traction type of loading; and traction loading is encountered in the draw works and to a certain extent in mud-pump operation. On the other hand, since a separately excited generator is used, a separately excited motor would be a wholly identical machine—certainly a desirable feature. With an added control system, the torque characteristics of a separately excited motor would be improved. The choice between the series motor and the separately excited motor is not clear until all of the factors involved are considered. In the judgment of some, the choice is clearly the series motor.

For a major portion of the load demand, the series motor is a self-regulating machine. Its speed-torque characteristics follow the load pattern; high starting torque at low speed, high speed at low torque. Its self-regulating characteristic requires modification near the no-load end of the operational sequence. At light load, its decelerating characteristics need improvement; and at mechanical failure of the load, the motor overspeeds.

The separately excited motor is inherently unable to regulate itself; it has very poor torque response and therefore a control system is required to provide the needed characteristics. In other words, the motor itself will not respond to load change, it must be adjusted to meet the new load requirements. Its decelerating characteristics and performance at light load are good. While it is not affected by a mechanical loss of load, it must be protected against an electric (field) failure, which would cause it to "run away."

It thus becomes a question of which motor requires less effort, equipment, and cost to compensate for its disadvantages. It has been found that the advantages of the series motor outweigh the advantages of the separately excited motor, and its disadvantages are easier to deal with than the disadvantages of the separately excited motor. To sum up, the following are the advantages and disadvantages of each type:

Series-motor advantages

1. Speed-torque characteristics suited to the needs of the draw works, mud pump, rotary table, and auxiliary and miscellaneous drives.
2. Standardization of all motor drives.
3. Simplified wiring for application.
4. Controls centralized in motor-control and generator-control cabinets.
5. Higher inherent torque characteristics.
6. Wider speed range.

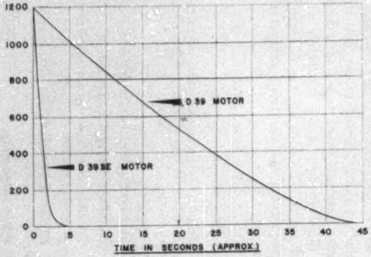


Fig. 4 (left). Typical draw-works deceleration curve for D39 and D39SE motors

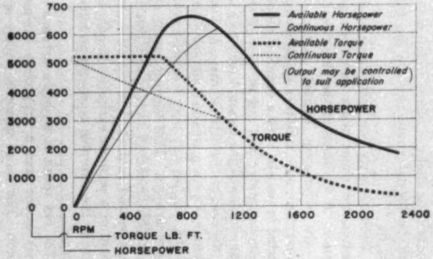


Fig. 5. D39 industrial motor, torque and horsepower curve for applications from 150 to 625 hp

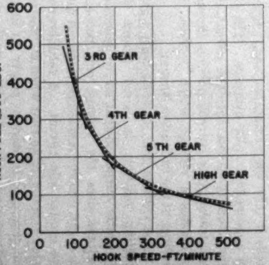


Fig. 6. Typical 1,200-hp draw-works curve compared with theoretical "ideal" curve

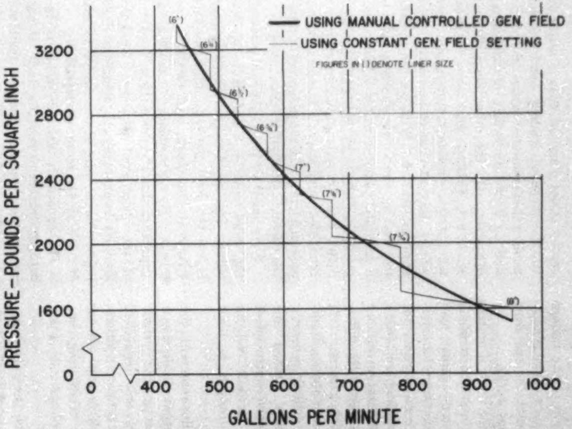
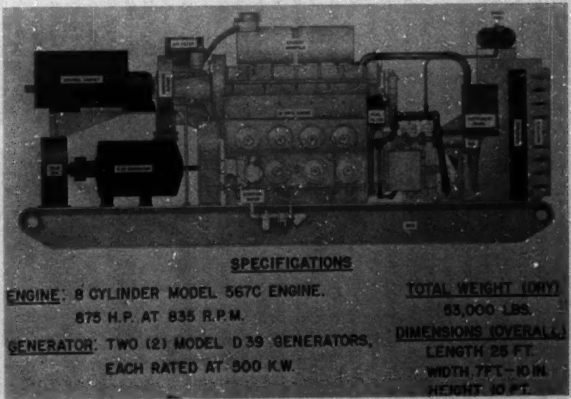


Fig. 7. Typical 1,000-hp mud-pump application curve using two D39M motors



SPECIFICATIONS

ENGINE: 8 CYLINDER MODEL 567C ENGINE.
875 H.P. AT 835 R.P.M.

GENERATOR: TWO (2) MODEL D 39 GENERATORS,
EACH RATED AT 500 K.W.

TOTAL WEIGHT (DRY)
53,000 LBS.

DIMENSIONS (OVERALL)
LENGTH 25 FT.
WIDTH 7 FT.–10 IN.
HEIGHT 10 FT.

Fig. 8. Unit of 875 hp

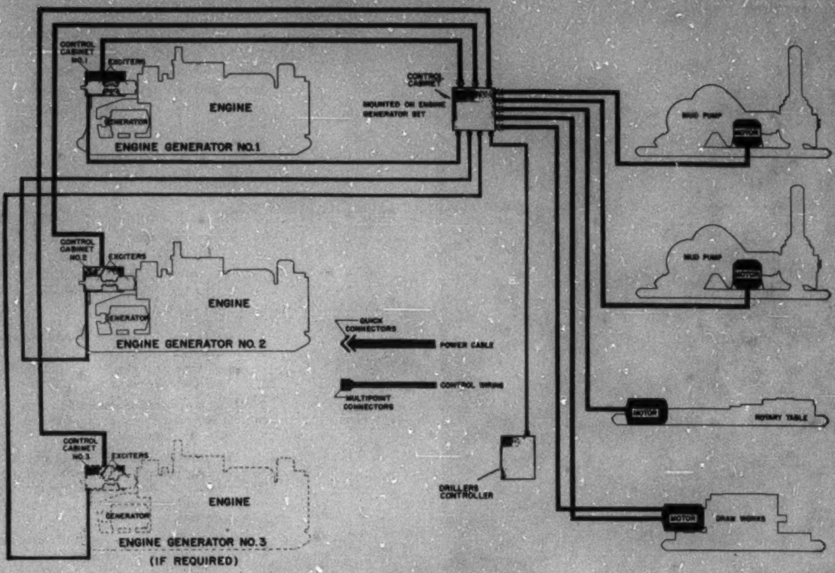

CONTROL CABINET NO.1
EXCITERS
ENGINE
GENERATOR
ENGINE GENERATOR NO.1
CONTROL CABINET
MOUNTED ON ENGINE GENERATOR SET
MUD PUMP
MOTOR
CONTROL CABINET NO.2
EXCITERS
ENGINE
GENERATOR
ENGINE GENERATOR NO. 2
MUD PUMP
MOTOR
QUICK CONNECTORS
POWER CABLE
CONTROL WIRING
MULTIPOINT CONNECTORS
MOTOR
ROTARY TABLE
CONTROL CABINET NO.3
EXCITERS
ENGINE
GENERATOR
ENGINE GENERATOR NO.3
(IF REQUIRED)
DRILLERS CONTROLLER
MOTOR
DRAW WORKS

Fig. 9. Diesel-electric drill-rig schematic

Series-motor disadvantages

1. Safety control for speed limit in case of load loss due to mechanical failure.
2. Deceleration control for performance improvement.

Separately excited motor advantages

1. Does not overspeed with loss of load.
2. Better inherent decelerating characteristics.

Separately excited motor disadvantages

1. Speed-torque characteristics do not match load requirements; the motor runs at constant speed and does not respond to change in drilling conditions.
2. Requires voltage-failure protection for each motor field.
3. Machine transients need to be controlled.
4. Three kw per motor is required for control power.

Motor-Overspeed Protection

Model *D39M* motor used in the electromotive diesel-electric unit is a d-c 4-pole series-wound machine, in which all parts, except the main-field windings, are interchangeable with the *D39* generators. It is conservatively rated at 625 hp for continuous operation. The machine has a continuous-current rating of 670 amperes with a short-time rating of 1,000 amperes for intermittent hoisting operations.

To protect the motor against overspeed, an *E*-type through-cable relay is used. It consists of four parts; a 2- by 3- by 1/4-inch magnet yoke, voltage coil, armature assembly, and contact assembly. The relay is biased by a generator current-carrying cable passing through the yoke of the relay, and operates on generator voltage by energizing a voltage-sensitive coil. In operation, the flux set up in the yoke by the two turns of the generator lead tends to keep the relay from picking up.

The relay operates to remove generator excitation when the generator voltage has reached the value which, for a given load, tends to result in motor overspeed. At this point the voltage-coil flux is of sufficient magnitude to overcome the biasing effect of the cable and to pick up the relay. Power to the motor is cut off until the driller shuts off his power control to re-establish the circuit.

The New S-E Motor Control

As can be seen from the previous description, one of the two shortcomings of the series motor, overspeeding in case of mechanical failure, is easily handled. The motor literally shuts itself off by the use of a simple, rugged unit.

Providing the desired degree of deceleration control for operating equipment such as the draw works takes some ingenuity. A device is needed that gives the series motor the better decelerating response of the separately excited motor to light load without losing any of the series-motor advantages or taking on any of the separately excited motor disadvantages.

The device proved to be the new "super-excitation" (S-E) control. In effect, it consists of a selenium rectifier and a step-down transformer supplied from the 220/-440-volt a-c utility system of the drilling rig. It maintains a small amount, approximately 20%, of excitation in the motor field.

At light or no load without the S-E control, the armature does not decelerate fast enough when power is removed. The S-E control slows the armature quickly and smoothly. It accomplishes this by converting the inertia of the motor and its load into electric energy which is in turn dissipated in the generator and engine. (The motor becomes a generator whose

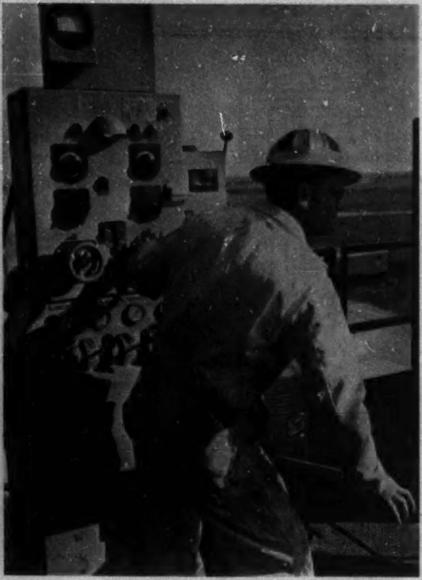


Fig. 10. Control stand

field is supplied by the superexcitation and drives the main generator and engine.)

So long as the normal excitation is at an appreciable value, the device has little if any effect on the motor characteristics. It becomes important only when the normal excitation falls off to the point where it is in the magnitude of the superexcitation current. The S-E control can be adjusted so that the decelerating current can be controlled to a desired value.

This device is a compact unit with no moving parts and with components that have long life. It can be added to a motor already in operation or built integrally into the control system.

Fig. 4 shows the effect of the S-E control on the operation of the electric transmission. Fig. 5 shows the torque and horsepower curves for the *D39M* motor. Fig. 6 shows the optimum draw-works curve and *D39M* operating curve for a typical draw-works application. The curve in Fig. 7 illustrates a typical application of two *D39M* motors to a mud pump rated at 1,000-hp input. Maximum pump output is shown by the saw-tooth curve and is obtained by adjusting the speed control knob to give full horsepower over the entire operating range of the pump.

Complete Unit

With the description of the load motors, the major d-c machines in the diesel-electric set have been covered. In summary, a typical power plant utilizing the previously described equipment would consist of the following:

1. One or more model *SR-8* skid-mounted engine-generator sets, shown in Fig. 8, each equipped with two dual-drive individually controlled d-c generators.
2. A separate skid-mounted motor-control cabinet which houses all motor-switching and protective equipment.
3. The driller's control panel containing all operating controls and alarm indicators in a compact cabinet which can be mounted integrally with the draw-works clutch and brake controls.
4. Model *D39* drive motors mounted integrally with the driven equipment.

A typical field arrangement of the power plant is shown in Fig. 9. In this case, dual-drive motors are used on two mud pumps and the draw works while a single motor drives the rotary table. The 8-position selector switch located on the driller's control stand (Fig. 10) selects the desired arrangement for coupling available generator output to the various motor-drive units. For full-torque low-horsepower operation, a single generator is connected to the individual motor drive. For full-torque full-horsepower operation, two generators are connected in series to the motor-drive unit.

◆

Discussion

G. W. Webb (General Electric Company, Dallas, Tex.): I wish to congratulate the author on a comprehensive coverage of this very timely subject. Early versions of diesel-electric rigs were complex and costly. It is perhaps fortunate for the drilling industry that locomotive market conditions permitted manufacturers of traction equipment to turn their attention to other applications at a time when offshore drilling became active.

We are in complete accord with the author's design criteria of:

1. Simplicity.
2. Interchangeability.
3. Flexibility.
4. Built-in protection.
5. Ease of operation and maintenance.

Means of accomplishing these objectives electrically is a vital concern of the industry, and I feel constitute a profitable area for study and discussion. The following comments are offered in the interest of furthering our mutual objectives.

With reference to generator selection, the author states that a differential-generator characteristic is required, but proposes a shunt-wound generator with an external current-regulating system to attain this characteristic. This was used in earlier versions of electric rigs.

True, a differential field in the generator requires space and could mean a larger machine; however, if the motors and generators are to be interchangeable in frame size and armatures, the choice seems to be between a reduction in generator peak rating and the extra controls required for current- and voltage-regulating systems.

Our experience leads us to question the use of such regulating controls on drilling rigs when the job can be done by a few turns of heavy copper in the generator field, which is simple to understand, the ultimate in reliability, requires no critical adjustments or maintenance, and is fast in transient response.

I would appreciate clarification of the objection to a differential generator for the reason that "multiplicity of load functions cannot be met without serious control complications." It appears that either type generator can be assigned to various loads on the rig with about the same number of switching contactors. Actually, it is quite simple to parallel two differential generators to obtain even higher intermittent output from the hoisting motors than could be obtained with a single higher rated shunt generator.

With reference to motor selection, it is felt that motor and generator selection should be considered together since it is their combined output which determines the significant performance characerisitics from the standpoint of driving the load.

In considering the relative advantages of the series motor, universally used on locomotives, and the separately excited shunt motor, which for years has been the accepted standard for electric drilling rigs, I would like to compare briefly some features of the locomotive- and drilling-rig applications.

1. The locomotive motor is geared to the wheel. Drilling motors are chained or belted.

2. Light loads or loss of load are difficult to obtain on a locomotive. Even so, series motors are occasionally destroyed despite years of development or overspeed devices.

3. There are no mechanical gear changes provided in a locomotive drive. Steam-type draw works, commonly used, have four mechanical ratios. Mud pumps have liner changes.

4. Line-pull, pump-pressure, and drill-stem torque limitations I believe have more serious consequences than wheel slip in a locomotive.

5. In a locomotive the main controls are located much closer to the motors than on a drilling rig. This permits the switching of power circuits to put the traction motors in series or parallel and to change their fields to obtain the long smooth tractive effort curves.

Now in considering the series-motor advantages for drilling applications, the following applies:

1. It is readily conceded that the inherent speed-torque characteristic of the series motor is closer to ideal performance. However, the differential generator in itself provides high speed at low torque, and low speed at high torque with the shunt motor. This combination inherently limits no-load speed without reliance on external control devices and also limits stalled torque without external current regulation.

2. With either motor all motors can be standardized.

3. Simplified wiring is assumed to refer to elimination of shunt field connections. It is questioned whether the necessity for reversing the series field on draw works and rotary, the superexcitation supply and leads, and the torque-limiting and overspeed protection would not defeat the purpose of simplified wiring.

4. It would appear that controls would necessarily be centralized for either motor.

5. Is it not true that the higher inherent torque of the series motor must be limited by external controls to the same permissible maximums.

6. The wider constant-horsepower speed range of the series motor is not questioned, but the difference without field switching on a 4-speed draw works would not seem to warrant the extra control complications.

The series-motor disadvantages are not questioned. We do, however, seriously question its use on any chained or belt-driven application where external overspeed shutdown is depended on to protect life and property.

With reference to the shunt-motor advantages, the shunt motor not only will not overspeed on loss of load, but since it is always excited at full field, its speed is always closely proportional to applied voltage, permitting acurate speed control at light loads, speed indication with standard voltmeters, and good regenerative braking characteristics. Shunt-motor torque is directly proportional to current as limited by the differential generator characteristic, and can be measured directly with a standard ammeter.

With reference to the shunt-motor disadvantages, the following applies:

1. Considering the combined motor and generator characteristic, the shunt motor will respond to changes in drilling conditions, since as the generator trades volts for amperes, the motor trades speed for torque or vice versa.

2. Thousands of shunt-motor applications in many industries would indicate that the field-failure hazard (which can happen only through field-circuit discontinuity simultaneous with loss of load) is not a serious problem. A simple voltage relay can be used to protect against this if required.

3. Transient-current surges which might cause difficulty are felt to be minimized on the limited engine-generator power supply and with the voltage-control systems normally used.

4. Field-excitation power of approximately the same magnitude is required for either motor, the series motor taking this power from the main engines and from the superexcitation source, while the shunt motor takes it from an exciter.

With regard to controls, many references have been made to controls in the preceding discussion. The 3-field differential-generator and shunt-motor combination requires only commonly familiar power and field contactors, simple relays, rheostats, and resistors. All rectifiers, saturable reactors, variable-voltage transformers, and current- or voltage-regulating systems can be eliminated. With the use of commonly accepted air throttles, pressure switches, and actuators, the generator-field rheostats can be remotely controlled, eliminating individual exciters for each generator. A single exciter with a stand-by is adequate for all rig requirements. It is felt that this approach is consistent with the objectives of simplicity, built-in protection, and ease of operation and maintenance without sacrificing interchangeability or flexibility.

L. L. Johnson (General Motors Corporation, La Grange, Ill.): Mr. Webb interprets the recommendation of the separately excited generator as an objection to a differential generator, and this is not intended. The result of carefully studying oil-rig load requirements revealed a very wide range of parameters. For example, consider the rotary table. Certainly it is desirable to limit table torque, and this limit will vary as some function of the drill pipe size. To provide such flexibility will require control and it certainly is more easily and more accurately accomplished by operating on a feedback control system than on the field of a differential machine. In the feedback system, power, speed, or torque can be varied independently or together to obtain almost any desired characteristic. In other words, a differential generator adjusted to give desirable draw-works performance does not fit either a rotary-table or mud-pump requirement. Compromising so that a single generator characteristic be usable for all loads sacrifices much performance; and to apply control to modify the differential machine requires about the same apparatus used basically in the feedback system.

Mr. Webb suggests paralleling generators for intermittent-high-intermittent output on a hoist—if motors and generators are identical, no one machine establishes the limit, they all do.

In considering the pros and cons of a motor to oil-well drill rigs, similarities to electric traction were logical, but not necessarily the details. There are particular problems peculiar to each. No attempt will be made here to compare locomotive wheel slip with draw-works line pull, or locomotive motor connection with mud-pump liner-size changes.

Proceeding to series-motor advantages, Mr. Webb cites the inherent no-load speed limit of separately excited motors; if this is true, why are field-failure relays provided? Either series or separately excited motors require a protective relay.

It is true that all motors can be standardized.

Simplified wiring does mean wiring to separately excited fields, and also the power supply for them. Only one or two loads on the rig require reversing, and to accomplish this only one additional cable between the load and control is required when the series motor is used. The superexcitation capacity amounts to only a small fraction of the normal motor excitation of the separately excited motors. It puts out nothing except when the pump back braking is being utilized, and there is no additional rig wiring to accomodate it.

The controls are centralized regardless of motor type, and the torque limit is provided by either basic generator characteristics control, or by control requested by the user. The accuracy of these limits is not subject to several variables, such as field temperatures and field supply voltages.

Mr. Webb's comment on shunt-motor advantages tends toward the negative by repeating previously stated problems in applying the series motors, and overemphasizing the assumed desirability of shunt-motor characteristics. The need for braking is not required and perhaps not desired for all loads. There are also many conditions where, to get the desired speed on a given load, either motor-field control or engine speeds much higher than power requirements justify (or both) are necessary and result in control complexity or excessive fuel consumption, or both. The fact is that the requirements can be, and are, satisfied by both types of motors.

With reference to the comment on disadvantage of separately excited motors, we find that appropriate machine response to conditions can, as stated, be achieved, but the degree of meeting all desired conditions is too involved to consider here.

Reference is made to thousands of industrial applications of shunt-excited machines; these thousands of shunt-wound motor applications can be reduced to a fraction of that number if only separately excited motors are considered. Excluding the special one-of-a-kind elaborate reversing drives and very accurate speed-regulated motors reduces the number still more. In addition to that, either separate or shunt excited are almost universally protected with field-failure relays; relays which are the equivalent of speed-limit relays employed where series motors are used. It should also be pointed out that in the case of separately excited machines total excitation failure is not the only hazard; the motor behaves as predicted only when the field current is as assumed for the calculation. There are severel prominent influences such as field temperature which affect the value of field current. Troubles in the field circuit of one motor can affect all others on the system unless the complication of protection control is applied, or a separate excitation supply provided for each motor. As a final comment on details of this nature it should be emphasized that there are probably many times as many points as brought up here that have been considered by those who have satisfactory performance in their assigned responsibility.

Finally, Mr. Webb mentions controls; control is equally essential regardless of motor choice, and detail for detail, there is not a great deal of difference in the number of items. Reference is made to a single exciter with stand-by being all that is required for a rig; surely there must be some wiring to apply the power from it to motor and generator fields, some control to permit transferring from one exciter to the other, some kind of protection to prevent a failure in one circuit from affecting all the others. Some "complications" are admitted such as remotely controlled rheostats. As a suggestion, there are severe limitations as to how remote the rheostats can be for a practical mechanical connection. There certainly are reasons why the world's largest and most successful supplier of rheostat-controlled 3-field generators did not choose that system for this application, and there is no time to treat them in detail here.

E. E. Hogwood (Westinghouse Electric Corporation, East Pittsburgh, Pa.): Mr. Hefner is to be congratulated on his timely paper describing General Motors' equipment for drilling rigs. Integrally designed diesel-electric drilling equipment have been available in past years from one of the major supply companies. These equipments utilized standard-oil-field diesel engines and industrial-type electric equipment. In most cases, the control was designed to meet the specific requirement of the drilling customer.

The advantages of the d-c variable-voltage oil-well-drilling transmission have long been recognized; however, their general usage has been limited due to the initial investment required when industrial apparatus tailored to the customer's requirements is used. The introduction of mass-produced transportation-type equipment with its lighter weight and reduced space requirements has made it possible to lower the first cost of a diesel-electric rig to be more competitive with mechanical drilling rigs.

With reference to load requirements, the author's analysis of the load requirements of a drilling rig are quite correct, although it should be recognized that the torque-versus-speed requirements of a locomotive are not exactly parallel to those of a drilling rig. First, the locomotive is basically used for accelerating a load and maintaining a particular speed for reasonably long periods of time. On the drilling rig, the load must also be accelerated and brought up to a speed, but the speed is not maintained for any appreciable length of time. Also, the draw works effectively contain a set of gears whereby the required constant-horsepower curve is better accomplished; therefore, it is not necessary for the motor to cover the entire speed-torque range. The greatest difference between the locomotive and the drilling rig is at light loads, such as "fishing" and "making up."

It is important to have relatively good speed control and the ability to accurately position a light load when performing these operations. For this reason we do not agree with the author that the series-wound motor adequately fits the draw-works and rotary-table applications.

With reference to motor characteristics, the separately excited shunt motor has been successfully applied on quite a number of diesel-electric drilling rigs, mine hoists, and cranes where an adjustable variable-voltage system is employed. The separately excited motor operating from a variable-voltage generator will regulate itself to the load in a manner similar to the series motor, although over not as wide a speed-torque range. The author has indicated that the shunt motor is inherently unable to regulate itself to the load but this is true only if the motor is operating from a constant potential bus. When a separately excited motor operates from a variable-voltage generator, i.e., one whose voltage output droops markedly with increased load, the motor will regulate itself quite well to the load. Its speed-torque curve approximates the desired constant-horsepower curve in the full-load region, and has the additional advantages of no-load speed limit and stalled torque limit without any external control devices. The sacrifice in speed-torque range is considered desirable to gain the advantages of limited no-load speed, limited torque, and inherent braking characteristics.

With regard to generator characteristics, a more direct way of obtaining the desired volt-ampere generator characteristic is to use a differential-series field in either the main generator or its exciter. The differential-field strength is proportional to the main-armature current. This method ensures an inherent drooping characteristic and consequent current and torque limit without the use of external regulators and other control devices. We have chosen the foregoing method of obtaining our drooping characteristics because we consider the inherent means more reliable and less complicated.

With reference to overspeed protection, a considerable number of separately excited d-c motors are presently operating on diesel-electric drilling rigs without loss of field protection. To our knowledge, there has not been one report of a motor running away. The heavy loads on the main drives prevent the motor from running away; therefore, loss of field protection is not generally required because simultaneous loss of field and load are quite remote.

In general, the case for the separately excited motor has been presented in this discussion; however, the characteristics of the series-connected motor are considered adequate for the mud-pump drives and other auxiliaries where speed control is not of too much importance. The foregoing comments pertain to the application of series- or separately excited shunt motors on the draw-works and rotary-table drives where reversing and good speed control are usually required. Mud pumps do not require this type of control.

B. H. Hefner: The author appreciates sincerely the interest evidenced in this paper and wishes to thank the discussers for their excellent comments. The discussions by Mr. Webb, Mr. Johnson, and Mr. Hogwood provide an opportunity to examine the various systems objectively and in detail.

We recognize the fact that in the design of electric machines there are several routes that can be followed to achieve an end result. Each system of rotating electric machines and associated control equipment can be justified on the premise that it accomplishes its design goal.

Admittedly, there are definite advantages and disadvantages to every power system. I shall not take issue with designs that differ from ours. Our choice of design was dictated by studies of a number of power systems used in a wide range of applications, under all conditions of field drilling.

In the interest of providing a constantly better product, engineers are striving continuously to develop and to improve the design of both machines and controls to obtain ever higher levels of performance and reliability.

Germanium Rectifier Equipment for Electrolytic Processes

L. G. MILLER
NONMEMBER AIEE

W. R. HODGSON
ASSOCIATE MEMBER AIEE

DIRECT-CURRENT power is being utilized along with alternating current in many industrial systems. Equipment employing sources of d-c power for electroplating, anodizing, electrochemical, adjustable speed drives, and other industrial applications are being used in increasing quantities. It is not convenient generally to power a combined a-c and d-c power system from two separate sources. Available power must either be rectified or inverted so that all electrical needs for a plant can be handled from the same source. The normal direction of change is from a-c to d-c, and it therefore seems reasonable that rectifiers will have considerable importance in the field of design and application for industrial systems.

There is always a demand to improve or develop electric equipment. Usually many years of tests, development, and field trials follow before a new development matures. Such is the case of germanium semiconductor rectifiers which are now coming into their own. The need for a low and medium voltage rectifying device spurred industry into a rapid development program to fill this gap. This energetic development of the germanium power rectifier has reached a point where its application is no longer restricted to the electronics and communications industry. Now the germanium power rectifier takes its place beside the other rectifying devices and promises to replace them in many applications.

Direct-current loads at almost any voltage and current can be supplied with germanium rectifier series-parallel combinations. The ultimate voltage and current size is generally a matter of economics. This paper will discuss germanium diodes, their construction, operation, and protection along with typical applications for germanium semiconductor rectifiers.

Description of the Semiconductor Diode Element

In order to compare this rectifier with other types, we should first become familiar with its physical construction. Fig. 1 shows a typical finished germanium diode. The heart of the semiconductor rectifier is a sandwich of five layers of material fused together. The central layer, the germanium itself, is a slice of single crystal n-type germanium. This single crystal is grown from a melt of ultrapure germanium to which small amounts of n-type impurities have been added. The top and bottom layers are the collectors which have approximately the same coefficient of thermal expansion as germanium, and have good heat and electrical conductivity. During the assembly process, germanium is soldered to the bottom collector layer with pure tin and to the top collector layer with pure indium. The bottom collector-tin layer is ohmic, but the indium on the top surface of the germanium forms a p-n junction by alloying into the germanium and changing its conductivity from n to p type in the upper region.

To form a workable rectifier device, the diode sandwich must have additional work performed on it. The diode assembly is soldered to the water-cooled copper base and a flexible lead is soldered to the top molybdenum layer. A glass-Kovar seal is then assembled around the diode and welded vacuum tight. The semiconductor rectifier is then tested for vacuum tightness with a helium leak detector. After degassing the internal parts, a dry gas is sealed in the semiconductor rectifier at atmospheric pressure. This hermetically sealed assembly assures that the diode surface will not be subjected to moisture and other impurities which may cause cell failures. The assembly must also be designed so that the crystal is not subjected to damaging mechanical forces, either from thermal expansion or outside forces such as shipping and handling.

A p-n junction has rectifying properties; these properties are illustrated in a simplified form in Fig. 2. If the p region is made positive with respect to the n region, holes from the p region and electrons from the n region are driven toward each other. The holes and electrons recombine, each electron filling one hole. This is the easy direction of current flow. If the polarity is reversed, the holes in the p region are attracted toward the negative polarity and the electrons in the n region are attracted toward the positive polarity, leaving what is essentially an insulator in the junction region and blocking the flow of current. Thus, this p-n junction allows current to pass unimpeded in one direction and offers a very high impedance to current flow in the opposite direction. Therefore the germanium rectifier described previously functions as a rectifier. Fig. 3 is a cutaway view showing the method of assembling a typical diode into a unit to provide cooling, hermetic sealing, and mechanical protection. Fig. 4 shows the completed assembly.

Method of Rectifier Protection

The germanium diode having such a small mass must have proper protection. With proper protection the unit can be relied upon for continuous unattended reliable operation. This protection may conveniently be divided into several features as follows:

1. Overload
2. Surge voltage
3. Junction failure
4. Cooling water failure

The semiconductor rectifier diode has a very short thermal time constant compared to that of transformers, reactors, buswork, and other pieces of equipment with relatively large masses. For this reason, it cannot be protected against overloads by ordinary thermal inverse time devices. In general then, overload protection will be provided for by means of overcurrent relays on each parallel rectifier together with standard devices to protect the transformer and other associated equipment.

A diode failure occurs whenever the cell loses its ability to block inverse voltage. This places a short circuit on the transformer windings. In single-way circuits it also places a short circuit across the d-c bus. This short circuit will cause fault current to flow in the good diodes. A fast acting current-limiting fuse in series with the faulted diode will interrupt this fault current before it has built up to destructive values in the good units. A fuse-monitoring system provides an indication when a fuse has blown. Each fuse has an indicator on the outside to determine

---

Paper 57-1178, recommended by the AIEE Industrial Power Rectifiers and Semiconductor Metallic Rectifiers Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference on Rectifiers in Industry, Chicago, Ill., June 4–5, 1957. Manuscript made available for printing September 20, 1957.

L. G. Miller and W. R. Hodgson are with Westinghouse Electric Corporation, East Pittsburgh, Pa.

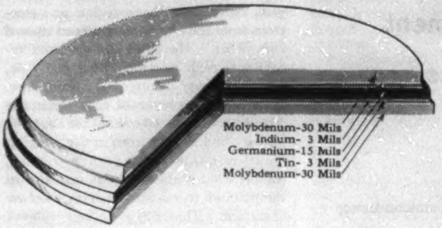


Fig. 1 (left). Cutaway diagram of a germanium rectifier cell

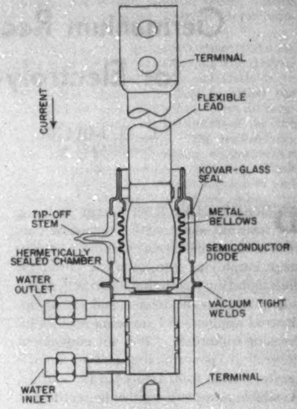


Fig. 3 (right). Water-cooled semitron rectifier cutaway view

which fuse has blown. Other means of removing the faults can be used depending on the number of units in parallel and the type of service.

The most common method of cooling is by liquid cooling which quickly removes heat. Thus the diodes are assured of the best thermal conditions for maximum life, and are independent of daily fluctuations of ambient air. Since the diodes are water cooled, means must be provided to protect against water supply failure. Consequently, if the water supply should fail, the rectifiers are shut down and an alarm is given.

Methods of Output Control

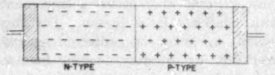
The semiconductor diode is a free conducting device. That is, it will start conducting whenever the *p*-type material is made positive with respect to the *n*-type material. For this reason, control of the output voltage must be obtained through varying the input voltage or by changing the regulation of the rectifier by means of a variable reactance. The most common method of varying the input voltage is with induction regulators or step regulators with no-load or load taps or both. The most common means of varying the reactance is by means of saturable core reactors. The choice of the foregoing types of voltage control depends upon the application requirements and the economics of the types available.

The induction regulator provides a continuous variation of the input voltage and maintains the inherent high power factor and low output voltage ripple of the rectifier even at reduced output. The size of the regulator is dependent on the range of control desired. It would be most suitable for applications requiring a steady output with small a-c line variations.

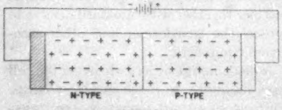
The step regulator can vary the input voltage either by load or no load taps or a combination of both. Although the output is not continuously variable over the range, it can be made so with steps of sufficient fineness for some applications. It can also take the form of manual tap switching of taps on the rectifier transformer. The step regulator also maintains a high power factor and low ripple at reduced output.

The saturable reactor is the other chief means of varying the rectifier output. This type of control functions by varying the reactance and, hence, the regulation of the rectifier. Full range, stepless control is obtained by varying the d-c bias current in the saturable reactors. This control is obtained with very little power and it is fast, lending itself readily to applications requiring automatic voltage or current regulators.

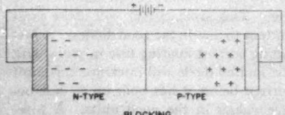
The saturable reactor provides a different output characteristic than the induction or step regulator. From about 0 to 90% output voltage, the output current is approximately constant for a given bias and is proportional to the bias. Above 90% voltage the regulation curves are similar to those with no saturable reactor. The no-load voltage varies very little with bias. Variations of supply voltage affect the no-load voltage but do not change the characteristics below about 90% of normal voltage. This means that below 90% of normal d-c voltage, the output current will not vary appreciably with load resistance changes or with input voltage variations for a fixed bias current. This constant-current output characteristic provides a cur-



JUNCTION DIODE



CONDUCTION



BLOCKING

Fig. 2. Semiconductor diode rectifier p-n regions

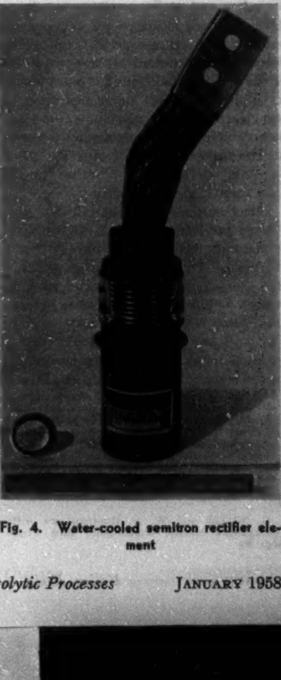


Fig. 4. Water-cooled semitron rectifier element

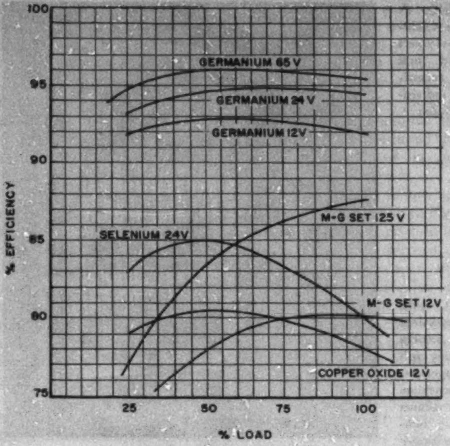


Fig. 5. Comparative efficiencies for several types of 4,000-ampere low-voltage conversion equipment

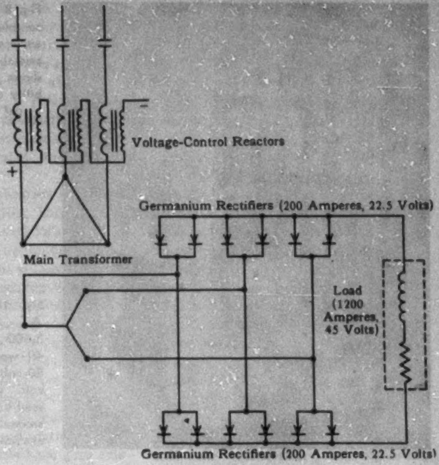

Voltage-Control Reactors
+
−
Germanium Rectifiers (200 Amperes, 22.5 Volts)
Main Transformer
Load (1200 Amperes, 45 Volts)
Germanium Rectifiers (200 Amperes, 22.5 Volts)

Fig. 7. A typical circuit for 6-phase double-way germanium rectifier

rent-limiting feature as well as a current-regulating characteristic. Since for a given output voltage the output current is very nearly proportional to the bias, a very simple voltage-compensation feature can be obtained by making the bias current proportional to the load current. This steep rise in voltage for very slight reductions of current provides arc stabilization for arc-casting furnaces.

A typical application of germanium semiconductor power supplies, with saturable core-reactor-type control is for continuous strip plating. The strip is continuously fed through several plating tanks with a rectifier for each plating tank. The total thickness of plating is directly proportional to the sum of all the rectifier plating currents and inversely proportional to the speed of the strip through the plating baths. Each rectifier has its own saturable core reactor. The d-c bias windings of all the rectifiers are connected in parallel to a regulated d-c bus. This regulated bus controls the total output of all the rectifiers in direct proportion to the line speed to maintain a uniform plating thickness with strip speed variations. The thickness of the plate can be varied by changing the regulating setting. The saturable core-reactor-type of control lends itself very well to this type of regulator application.

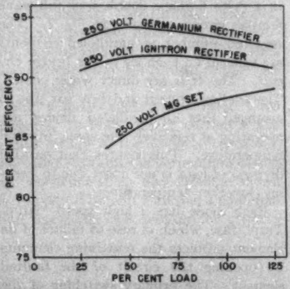


Fig. 6. Comparative efficiencies for germanium rectifier, Ignitron rectifier, and mg sets at 250 volts

Comparison of Germanium Rectifiers with Other Types of Existing Rectifiers

The germanium semiconductor rectifiers when compared to other metallic rectifiers or to Ignitron rectifiers have certain inherent advantages. One of the most prominent advantages of germanium is its efficiency advantages. Reference is made to Fig. 5 showing the relative efficiencies of germanium along with copper oxide, selenium, and motor-generator (mg) sets. The curve depicts very clearly the large efficiency advantage over the other conversion methods in this low-voltage class.

Fig. 6 is an efficiency comparison of a 250-volt semiconductor rectifier with a comparably rated Ignitron rectifier and mg set. A large efficiency advantage in favor of the germanium is evident. The high arc drop of the Ignitron rectifier compared to the low forward drop of the germanium rectifier, even with the required number of diodes in series, provides the efficiency difference.

These large efficiency advantages provide a real savings in dollars to the ultimate user particularly where large blocks of power are to be utilized. The cells can be operated at very high current densities due to the low losses within the cell. It is these low losses that are responsible for the germanium rectifiers high efficiency. This saves dollars in copper, cooling equipment, and power costs; this also provides additional savings in size and weight of a germanium rectifier. These rectifiers do not have a forward aging characteristic as do selenium or copper oxide. It is therefore not necessary to adjust transformer taps over a period of time to compensate for changing characteristics. Thus there will be no reduction in initial efficiency as elapsed operating time increases.

A considerable savings in space, installation cost, and original building cost can be experienced with germanium semiconductor rectifiers. The high current density that germanium can be worked at in normal operation when compared to other metallic rectifiers provides real space saving. This is due to the fact that germanium can be worked at

Fig. 8 (left). A complete 4,000-ampere 12-volt assembled unit shows accessibility of individual diodes

Fig. 10 (right). A complete 3,000 - ampere 40 - working - volt 80-volt open-circuit germanium semi-tron vacuum arc-melting furnace power supply

approximately 1,000 times the current density of selenium for the same forward drop, and approximately 300 times the current density of copper oxide for the same forward drop.

The germanium rectifier being a static device with no rotating parts means a savings in maintenance. All the component parts are designed and manufactured for many years of operation. No auxiliary circuits are required with semiconductor rectifiers as on mercury arc rectifiers with the resulting space and maintenance savings. Corrosion problems are minimized on the equipment and eliminated on the diodes due to their hermetically sealed construction.

Low installation costs are inherent with these devices since no special foundations are required. They can be quickly installed as it is generally a matter of making a-c and d-c connections along with the necessary water connections. Individual units can be arranged for close coupling when series or parallel connections are required.

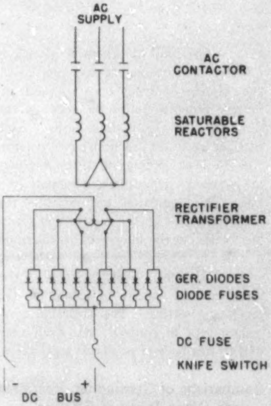


Fig. 9. A typical circuit for 6-phase single-way germanium rectifier

Industrial Application

The field of germanium rectifier applications is broad, and it encompasses low-voltage plating and anodizing supplies up to higher voltage electrochemical and industrial applications. A few typical installations will be discussed in this paper.

Recently a large copper company installed additional refining capacity and the decision was made to purchase germanium rectifiers after thorough investigation of all types that are available in this voltage class. The equipment is utilized in electrolytic refining of copper. The load is constant at 16,000 amperes with a potentiality of 135 volts direct current. At present the equipment is capable of 16,000-ampere output at 90 volts direct current.

A group of four 4,000-ampere packages were installed along with their switching and transforming equipment. Each 4,000-ampere unit is composed of a 6-phase full-wave double-way bridge circuit with two such circuits connected in series for 90 volts direct current. This circuit connection is shown in Fig. 7. A throat-connected Inerteen-insulated self-cooled rectifier transformer provides the proper element voltage. In the same tank as the rectifier transformer is a saturable core reactor to provide current and voltage control of the units. The primary switching of the rectifier units is by means of 2,400-volt oil-immersed contactors.

The units are controlled from a central control panel which includes all metering and control switches conveniently located in one position. The station operates unattended except for periodic meter reading trips by the cell room personnel. The germanium cells themselves are arranged in a delta 6-phase double-way full-wave rectifier circuit as shown in Fig. 7. The required number of paralleled elements are utilized on each leg of the bridge circuit. The cells are direct water cooled. The cooling system and bus bar are so arranged that all paralleled diodes are operating at relatively the same junction temperature. This assures that parallel division of load is not upset due to varying junction temperatures.

Each diode has a high speed "Amp-Trap" fuse, which in case of failure of an element protects the remaining elements by opening the circuit of the faulted element. The primary switching of the rectifier transformers is by means of motor starter-type high-voltage oil-immersed contactors. These contactors are bussed together to form a co-ordinated line-up with the a-c switchgear and d-c control compartment. Overload protection of the individual rectifier units is by means of high-speed relaying from the d-c side of the unit. Transformer fault protection is provided in the primary contactors by means of disconnecting-type fuses similar to those on motor-feeder circuits. An incoming line auxiliary compartment together with the station auxiliary transformer and the a-c incoming line breaker make up the remaining complement of a-c metal-clad equipment. Close coupled to the a-c switching equipment is located a d-c control and metering cubicle. All individual unit contactor switches, d-c ammeters, and individual and master manual current-control switches are conveniently located for the operator.

The rectifier transformer and saturable core reactor are in the same tank and are Inerteen-insulated, self-cooled, and throat-connected to the germanium rectifier. The transformer is connected delta on the primary and triple delta on the secondary. Each secondary delta is capable of furnishing 4,000 amperes at 45 volts direct current. The windings are so arranged that the rectifiers may be connected in series for 45, 90, or 135 volts direct current.

Included in the same tank is a saturable core reactor to provide manual current control for each transformer–rectifier combination. An additional feature of this saturable core reactor is the current limiting action of the reactor under short-circuit conditions. This reactor also to a certain degree limits inrush currents on the unit.

The foregoing application depicts a large installation where transformers, metering equipment, and switching equipment are not an integral part of the rectifier cubicles. For some applications such as tin-plating, arc-furnace supplies, anodizing, cathodic treatment, and cleaning applications the power level is such that a completely self-contained unit can be supplied.

Fig. 8 shows such a unit rated 4,000 amperes, 12 volts, direct current which would fit any of the above applications. This unit contains the germanium diodes connected in a single-way circuit as shown in Fig. 9. The unit contains a combination linestarter–AB breaker for primary protection and switching. A saturable core reactor is utilized for manual voltage control and a dry-type rectifier transformer, and germanium water-cooled diodes complete the assembly along with the necessary metering. This unit can easily be adapted to automatic current control as required on large tin plate lines and operated in parallel with many other like units.

A rather unique application is that of the unit shown in Fig. 10. This is a self-contained unit that is used for a vacuum arc-melting furnace. It is unique in that it has a high open-circuit voltage to strike the arc and a low operating voltage to maintain the arc. Operation of this unit at short-circuit conditions is not uncommon. The unit is rated 3,000 amperes at 40 working volts and 80 volts open circuit. Manual control is provided to preset the operating point and the furnace control then takes over to maintain the correct arc voltage and current. Recently four of these units have gone into service.

A conservative working level in peak inverse voltage of germanium is approximately 90 volts, therefore, a limitation on d-c voltages exists from a single element. As applications of higher d-c voltages are required, elements are placed in series. A typical example is a 300-kw 250-volt d-c industrial unit for crane power supply. A double way circuit similar to that shown in Fig. 7 is utilized except series elements are utilized to obtain the maximum d-c voltage required for this application. This unit station layout is similar to an Ignitron unit substation, that is, a metal-clad line-up consisting of an a-c breaker cubicle, a transformer cubicle, a germanium rectifier cubicle, and a control and d-c breaker cubicle.

A Magamp voltage regulator on this unit provides essentially flat voltage from no load to full load and allows compensating circuits to be used for proper division of load during parallel operation. This unit can be applied in existing substations where Ignitrons and mg sets already exist. The large efficiency advantage of the germanium rectifier at this voltage is shown in Fig. 6. Along with this efficiency advantage, low installation costs and relatively little maintenance make the semiconductor rectifier attractive in this field.

The afore-mentioned applications represent low-voltage units with a single diode in a single-way circuit or a series of diodes for higher voltages. For voltages in the 150-volt range, a different scheme is utilized. Here it would be necessary to put two diodes in series which means that if one diode fails, the peak inverse voltage applied to the second diode in series is doubled. The possibility of this diode failing is high. Therefore, in this voltage range, multiple-winding transformers are utilized and the rectifier units connected in series for the higher voltage. Two of the circuits shown in Fig. 7 are connected in series. The transformer therefore determines the peak inverse voltage that a diode will be subject to and failure of a single diode does not stress other diodes in the circuit.

A typical application in this category is the refining of cobalt which requires 150 volts direct current and amperes in the range of 8,000. At this power level, and due to the number of diodes in a unit, the transformer and regulating equipment would generally be separately mounted. A typical application consists of two 4,000-ampere packages consisting of an outdoor transformer, a rectifier cubicle, necessary a-c and d-c switching, and a control cubicle.

In addition to applications described in the foregoing, future applications in the 300-volt electrochemical field appear good; applications for hydrogen gas production, anodizing, and low voltage electroplating are also prominent.

Conclusions

Although a relatively new device, the germanium rectifier has already become well established. The device itself is designed to obtain long life, low maintenance, high efficiency, and good service reliability. It is demonstrating its performance in electroplating, anodizing, electrolytic processes, and other related applications. Operating experience to date with germanium rectifiers has been good. The future appears bright for the germanium rectifier, evidenced by the activity along this line now taking place in industry.

Reference

1. POWER RECTIFICATION VIA GERMANIUM AND SILICON, J. L. Boyer. *Westinghouse Engineer*, Pittsburgh, Pa., Sept. 1954, pp. 183–86.

A High-Current-Density Selenium Rectifier

C. C. GEIB
NONMEMBER AIEE

W. E. BROWN
NONMEMBER AIEE

THE semiconductor properties of selenium were discovered by Fritts in 1883. However, it was not until the period between 1939 and 1950 that the selenium rectifier came into its own in the United States as a means for converting relatively large amounts of power from alternating to direct current.

Up to the present, the development of the selenium rectifier has been characterized by a steady increase in the reverse voltage rating. The earliest voltage rating per cell was in the neighborhood of 12 volts, as compared to the most recent rating of about 45 volts. While the voltage rating was being increased, the current rating remained at approximately 160 milliamperes per square inch of rectifying area. Failure to increase the current rating may be explained by the fact that the current density of the selenium rectifier had a direct bearing upon the aging of the rectifier. Aging is defined here as an increase in the forward resistance of the rectifier and the resultant decrease in the output voltage.

The conventional selenium rectifier, when operated at the accepted current density of 160 milliamperes per square inch of rectifying area will, in general, reach a somewhat decreased but fairly stable output level within 20,000 hours of continuous service. After this initial decrease, there is an extended period of operation during which there is relatively little change in output. Rectifiers operating at less than normal current loads or on intermittent duty applications show a longer aging time. In view of this it has been assumed that an ampere-hour law governs the initial rate of rectifier aging. Conventional rectifiers, operating at current loads in excess of the accepted standard normal, follow this law to a great extent. However, a current density point is reached above which rather rapid and continuous aging takes place, even though the rectifier is maintained at a reasonable temperature through forced cooling.

The authors have been instrumental in recent developments in selenium rectifiers which have resulted in the production of a cell having a current-carrying capacity of 265 milliamperes per square inch of rectifying area with ordinary convection cooling. This cell is unlike the conventional selenium cell in many respects; actually, it behaves more like the high-temperature selenium rectifier. It was partially evolved from development work which was aimed at the improvement of the high-current-density cell is 105 C (degrees centigrade). This similarity to the high-temperature cell accounts for the fact that in this new type of cell current density has far less effect upon aging. Consequently, with adequate forced cooling, it can accommodate continuous current loads of 1.18 amperes per square inch The voltage rating of these rectifiers is 15 volts rms per cell.

The data shown in the figures were compiled from tests conducted with rectifiers of the type shown in Fig. 1, the cells in this type of rectifier are $4^{1}/_{8}$ inches square, and they have an effective rectifying area of 14.2 square inches.

Fig. 2 shows typical aging characteristics of the high-current-density rectifier operating at a current density of 850 milliamperes per square inch. In addition to showing the aging characteristics, this curve also substantiates the ampere-hour law governing the initial rate of rectifier aging, for in this case the rate of aging decreases markedly after about 3,500 hours.

Figs. 3 and 4 illustrate the static characteristics of the high-current-density rectifier at ambient temperatures of 25 and 105 C. The nonlinear characteristics of the forward voltage drop curve give an indication of the regulation of the rectifier when it is operated at different loads.

Fig. 5 shows a curve of nominal output voltage versus current for a 3-phase bridge rectifier made with six cells operating into a resistance load. The alternating voltage was held constant at 15 volts during the test. As was to be expected, the maximum change in output voltage fell within the region of the lighter loads. The absence of a greater flexure of the curve, especially at lighter loads, can be accounted for by the following: As the current load is decreased, a reduction in rectifier cell temperature occurs. The forward voltage of the rectifier cells increases slightly as the cells become cooler. The higher forward voltage causes a decrease in output voltage. The lower output voltage tends to neutralize the sharper rise in output which normally accompanies lighter current loads.

With ordinary convection cooling, the high-current-density cell has a current-carrying capacity of 265 milliamperes per square inch of rectifying area. This rating is fixed for rectifier stacks constructed with the cells normally spaced. For rectifier stacks constructed with one to four cells, or those having wide spaced cells, the convection-cooled rating is 345 milliamperes per square inch of rectifying area. Normal spacing, as applied here, varies with the size of the cell. In the case of cells of $1^{3}/_{4}$ inches in diameter, normal spacing is approximately 0.156 inch, center to center. Normal spacing for the 5- by 6-inch cells is about 0.312 inch. Wide spacing is approximately 1.6 times normal spacing. Table I lists some of the rectifier cell sizes available and their ratings with convection cooling.

To obtain the maximum output of 1.18 amperes per square inch from the high-current-density rectifier, adequate cooling must be maintained. Adequate cooling in this case refers to any method which will maintain a cell temperature below 105 C. The most common and effective method used to obtain the desired results is to employ a forced air draft. It has been found, during tests, that in order to obtain the maximum cooling effect with a minimum forced draft, a number of good practices should be observed. First, to avoid confusion, all measurements of air flow are made on the leeward side of the rectifier. Several measurements are made between the cells so that an average air flow through the rectifier can be determined. Air flow should be parallel to the plane of the rectifier cells. Conditions causing air turbulence such as obstructions which might impede the flow of air are to be avoided. Finally, wherever practical, the rectifier should be mounted, with relation to the air flow so that the air passes through the smallest dimension of the rectifier. For example, optimum cooling would be obtained in a rectifier constructed of cells which are 5 inches wide and 6 inches long if the air is made to enter and leave a path just 5 inches long instead of 6 inches long.

Paper 57-1174, recommended by the AIEE Industrial Power Rectifiers and Semiconductor Metallic Rectifiers Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference on Rectifiers in Industry, Chicago, Ill., June 4–5, 1957. Made available for printing September 20, 1957.

C. C. Geib and W. E. Brown are with the Fansteel Metallurgical Company, North Chicago, Ill.

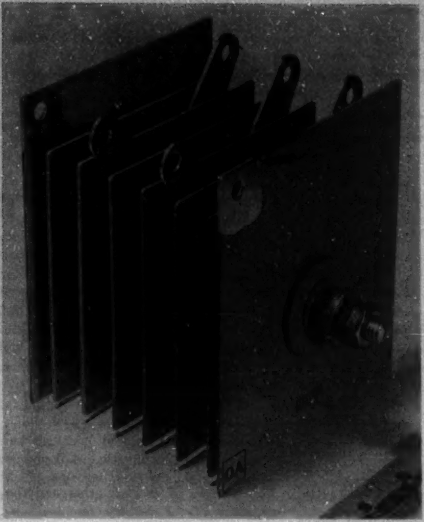


Fig. 1. Three-phase bridge rectifier made of $4\frac{1}{2}$-inch square cells; current rating 50 amperes

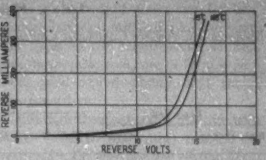


Fig. 4. Static characteristics, reverse direction

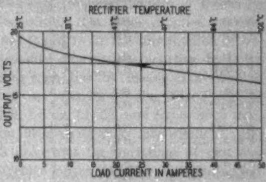


Fig. 5. Regulation curve: 3-phase bridge rectifier

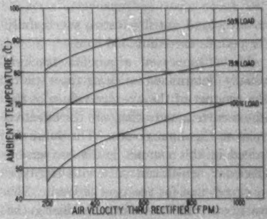


Fig. 6. Forced cooling curves: current loads based on density of 1.18 amperes per square inch

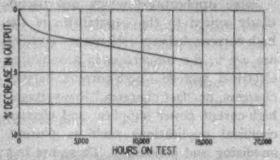
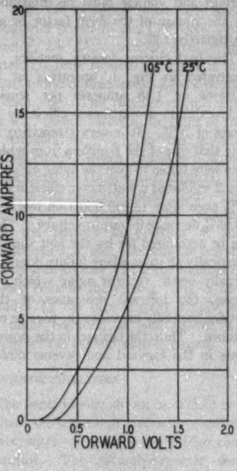


Fig. 2. Typical aging characteristics with current density of 850 milliamperes per square inch

Fig. 3 (below). Static characteristics, forward direction



In the design of a rectifier stack for applications requiring forced cooling, a maximum cell spacing of about 0.6 inch is employed, provided that such spacing is consistent with other aspects of the design such as physical dimensions, circuit connections, and type of load. The rectifier terminals, which have a considerable effect on cooling since they are a cause for air turbulence, are made of a material as thin as is practical for a particular design.

Fig. 6 shows the air velocity required, at various ambient temperatures, to maintain the temperatures of the rectifier cells under 105 C. The per-cent load on this graph is based on a current density of 1.18 amperes per square inch.

Fig. 1 illustrates a typical high-current-density rectifier. This particular stack is referrred to as an automotive rectifier and, as its name implies, is one of several stacks designed especially for a-c to d-c conversion in automotive electric systems. This is a 3-phase bridge rectifier which has a current rating of 50 amperes d-c with forced cooling. The maximum a-c input rating is 15 volts rms.

The automotive application presents a particularly rigorous test of both the physical and the electrical capabilities of a rectifier. The environmental conditions which the rectifier must be able to withstand include rather severe vibration, in addition to corrosive atmospheres. Some of the materials with which an automotive rectifier might be expected to come in contact, either directly or in vapor form, are gasoline, oil, alcohol, ethylene glycol, road salts, dust- and grit-laden air, and moisture. Obviously, in order to function properly for a satisfactory length of time, the rectifier must be protected from these environmental conditions.

Rectifiers designed to withstand extreme environmental conditions are protected by three coats of an epoxy resin base paint. Improved methods of application and curing of this finish make possible a 50-hour guarantee against a salt fog corrosion test of the type described in the military standard *202*.[1] Rectifiers having the protective finish just described have been able to withstand salt fog testing for periods in excess of 100 hours with no indication of corrosive attack. On occasion, finished rectifiers have been able

Table I. High-Current-Density Selenium Rectifier Cells, Standard Sizes and Ratings

Maximum D-C Ampere Continuous Output of Element with One Cell Per Arm at Ambient Temperature 113 Degrees Fahrenheit (45 C) or Less
Maximum A-C Volts Per Cell, 15

Basic Cell		Single-Phase		Three-Phase			Maximum Continuous D-C Rating as Valve, 113		Maximum No. of Cells Per Stack	Approximate Cell Spacing On Center
Size, Inches	Type	Half-Wave	Bridge or Center Tap	Half-Wave	Bridge	Center Tap	Amperes	Blocking Voltage		
$2^5/_8$ D*	*HE*	1.270	2.54	3.4	3.810	4.6	1.95	12.5	1 to 4	
		0.980	1.95	2.6	2.940	3.6	1.45	12.5	5 to 40	0.250
$2^5/_8$ D	*HE*	1.270	2.54	3.4	3.810	4.6	1.95	12.5	1 to 28	0.390
$3^3/_8$ D	*HF*	2.6	5.2	6.6	7.8	9.4	3.9	12.5	1 to 4	
		1.950	3.9	5.2	5.850	7.3	2.9	12.5	5 to 40	0.250
$3^3/_8$ D	*HF*	2.6	5.2	6.6	7.8	9.4	3.9	12.5	1 to 28	0.390
$4^3/_8$ D	*HG*	4.2	8.4	11.5	12.6	15.8	6.5	12.5	1 to 4	
		3.25	6.5	8.6	9.75	12.2	5.0	12.5	5 to 40	0.250
$4^3/_8$ D	*HG*	4.2	8.4	11.5	12.6	15.8	6.5	12.5	1 to 28	0.406
$4^1/_8$ Sq†	*HT*	4.9	9.8	13.1	14.7	18.2	6.5	12.5	1 to 4	
		3.75	7.5	9.9	11.2	14.0	5.0	12.5	5 to 40	0.250
$4^1/_8$ Sq	*HT*	4.9	9.8	13.1	14.7	18.2	6.5	12.5	1 to 28	0.406
$4^1/_8$ by 6	*HP*	7.3	14.6	19.5	21.9	27.7	11.8	12.5	1 to 4	
		5.7	11.4	15.3	17.1	21.2	8.4	12.5	5 to 40	0.312
$4^1/_8$ by 6	*HP*	7.3	14.6	19.5	21.9	27.7	11.8	12.5	1 to 28	0.500
5 by 5	*HH*	7.3	14.6	19.5	21.9	27.7	11.8	12.5	1 to 4	
		5.7	11.4	15.3	17.1	21.2	8.4	12.5	5 to 40	0.312
5 by 5	*HH*	7.3	14.6	19.5	21.9	27.7	11.8	12.5	1 to 28	0.500
5 by 6	*HR*	9.1	18.2	24.4	27.3	34.0	13.5	12.5	1 to 4	
		7.0	14.0	18.7	21.0	26.1	10.4	12.5	5 to 40	0.312
5 by 6	*HR*	9.1	18.2	24.4	27.3	34.0	13.5	12.5	1 to 28	0.500

* In diameter.
† Square.

to withstand salt fog testing successfully for periods exceeding 500 hours.

The efficiency of a rectifier stack is usually determined as being the ratio of the product of the average values of the direct current and voltage to the effective value of the input power. The value obtained in this manner is the conversion efficiency, which can vary with the type of circuit and for the type of load. The wattmeter efficiency of the rectifier is the ratio of the d-c power output to the a-c power input, both values being measured by means of wattmeters. This is the true efficiency of the rectifier, since it is dependent upon the effective values of input and output power. In general, the difference between conversion efficiency and wattmeter efficiency will be small in the case of 3-phase circuits and single-phase circuits with capacitance loads. However, in order to change from conversion efficiency to wattmeter efficiency in the case of a single-phase bridge rectifier operating into a resistance load, the product of the average values of the direct current and voltage must be multiplied by the square of the form factor or approximately 1.23.

The high-current-density rectifier, as illustrated in Fig. 1, operating at 50 amperes or 1.18 amperes per square inch, exhibits a wattmeter efficiency in excess of 70%. It is very interesting to note that one of the rectifiers from which data were taken to give the aging curve of Fig. 2 exhibited a wattmeter efficiency of 70% even after 13,000 hours on test and a 3.95% decrease in output voltage. This can be accounted for by the fact that in practically all cases where failure does not actually occur, rectifier aging which increases the forward resistance of the rectifier normally increases the reverse resistance. Thus, the changes in the power losses in the forward and reverse direction tend to offset each other and the efficiency of the rectifier remains substantially the same.

Some applications which are particularly suited to the capabilities of the high-current-density selenium rectifier are a-c to d-c conversion in automotive electrical systems, high-current battery chargers, or "fast chargers," low-voltage–high-curent power supplies, and electrochemical applications such as plating, anodizing and polishing. These are but a few of the many fields of applications open to this type of selenium rectifier which has the ability to carry unusually large amounts of current without the rapid deterioration of the rectifying characteristics.

Reference

1. Test Methods for Electronic and Electrical Component Parts. Military standard *MIL-202A*, Washington, D.C., Oct. 24, 1956.

A 10-Kw Germanium Rectifier for Automatic Power Plants

E. A. HAKE
ASSOCIATE MEMBER AIEE

THE AVAILABILITY of high-current semiconductor rectifying devices, high-grain magnetic amplifiers, transistors, and semiconductor diodes prompted this development. The practical objective of this design was to develop a rectifier of 10-kw rating that could compete with a motor-generator set and its controls for use by the Bell System. Higher efficiency and a reduction in over-all size over rectifiers of earlier design were the major inducements.

The design of the saturable reactors was the chief problem, where cost and reproducibility were hurdles to be cleared. Also, protection, maintenance, and versatility of the unit had to be explored, as related to its use in past, present, and proposed automatic power plants. A preliminary investigation was made to determine the feasibility of using square-loop saturable reactors controlled directly by a transistor amplifier with two 2-watt transistors in the output stage. This fundamental approach is described in a previous paper.[1]

Circuit Description

GENERAL

Fig. 1 is a block diagram of the rectifier which is rated at 50-volts 200-amperes continuous duty and 63-volt 200-amperes intermittent duty operating into a battery and resistance load. The power section of the rectifier consists of a self-saturated magnetic amplifier, with six germanium fan-cooled power rectifiers in a 3-phase full-wave bridge configuration. The control is by means of a transistorized push-pull amplifier using germanium transistors and rectifying diodes; and silicon diodes for blocking and as a voltage reference. Under normal conditions the rectifier operates as a constant voltage device up to rated load current. Slightly beyond rated load, the rectifier is automatically current-limited to prevent overloading.

The rectifier was designed to float a storage battery and to maintain its voltage to better than ±1%, and to operate at a constant current of 101–103% of rated output current for any condition of battery discharge down to 1.75 volts per cell. Taps are provided on the main transformer to furnish voltage for charging two or four additional cells which are a part of the reserve power facilities in a central office. Built-in test features are provided to simplify maintenance and adjustment.

POWER SECTION

The power section, shown in Fig. 2, consists of a 3-phase step-down transformer *T*1 furnishing alternating voltage to a self-saturated magnetic amplifier *L*1, *L*2, and *L*3 equipped with fan-cooled germanium power diodes *CR*1-*CR*6. An inductor, *L*4, *L*5, capacitor, *C*1-*C*10, inductor, *L*6, filter reduces the ripple in the output to the desired level; *L*4 is a swinging inductor. The rectifier output current lead passes through the cores of the d-c transformer, *CT*1, in the current sensing circuit, which provides a signal proportional to the d-c load current. A fuse is in the negative output lead. The switch *S*1 disconnects the rectifier from the battery in position 1, and in positions 2, 3, and 4 selects the correct taps on transformer *T*1 for the number of battery cells in the circuit. Fig. 2 also includes *T*2 the bias current supply for the main reactors, as well as the a-c supply for the current sensing circuit.

MAGNETIC AMPLIFIER

The basic circuit shown on Fig. 2 is a 3-phase self-saturated magnetic-amplifier power stage in a full-wave bridge configuration. The saturable-reactor control windings are driven by an amplifier having two 2-watt transistors in the output stage in a push-pull arrangement. The high power gain required indicated the need for cores using square-loop material. However, the selection of saturable-reactors for this application was governed by cost as much as the technical considerations. The saturable-reactor design selected was a compromise of these factors. Three-legged saturable reactors composed of two cores made of square-loop *DU* laminations having common d-c control windings satisfied the requirements of high gain, reasonable cost, use of form-wound coils, good reproducibility, and simple test procedures.

Although these saturable reactors were better balanced than other reactors investigated, the need for low output ripple and the limited margin of current overload of the germanium power rectifying elements, required additional means of balancing. The most effective method was found to be the introduction of adjustable amounts of negative feedback into each core by means of a rheostat across a separate winding (Fig. 2 winding 5-6, 7-8 on *L*1, *L*2, and *L*3) on each core, on the same leg as the gate winding. Fig. 3 illustrates the variations of rms and average current in each gate winding and power rectifying element with and without the balancing adjustment at full rated load current.

Unbalance causes high 60-cycle and 120-cycle components of ripple which the filter attenuates to a lesser degree than the 360 cycles for which it was designed. The output ripple measured at the battery was found to vary from 80 to 500 millivolts depending on the degree of unbalance of the various sets of reactors tested. Use of the balancing winding and its associated rheostat reduced the total ripple by a factor of 10 to 1. This adjustment will be made at the factory on the assembled rectifier. The process is simple, requiring only an oscilloscope connected to the input to the main filter and an a-c vacuum-tube voltmeter.

CONTROL SECTION

The control circuit is shown in Fig. 4. It consists essentially of a gate circuit to discriminate between voltage and current signals, and an amplifier operating in push-pull to drive the saturable-reactor control windings.

*CR*18 and *CR*19 are the gating diodes. Depending on the magnitude of the rectifier output current, a voltage signal received from the *ADJ VOLTS* potentiometer or a current signal from the current sensing circuit is connected to the base of transistor *Q*1 and is matched against the potential of the voltage reference diodes[2] *CR*13, 14, 15, 16 connected to the base of *Q*2. The amplified difference is applied to identical 2-stage current amplifiers consisting of *Q*3, *Q*5 and *Q*4, *Q*6 transistors, and associated apparatus. The push-pull output is connected to the control windings 11-12, and 13-14 of the saturable-reactors *L*1, *L*2, *L*3 shown in Fig. 2 through leads *A* and *B*.

Paper 57-1175, recommended by the AIEE Industrial Power Rectifiers and Semiconductor Metallic Rectifiers Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference on Rectifiers in Industry, Chicago, Ill., June 4-5, 1957. Manuscript made available for printing September 20, 1957.

E. A. HAKE is with Bell Telephone Laboratories Inc., New York, N. Y.

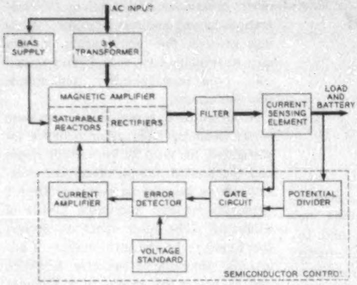

AC INPUT
BIAS SUPPLY
3φ TRANSFORMER
MAGNETIC AMPLIFIER
SATURABLE REACTORS
RECTIFIERS
FILTER
CURRENT SENSING ELEMENT
LOAD AND BATTERY
CURRENT AMPLIFIER
ERROR DETECTOR
GATE CIRCUIT
POTENTIAL DIVIDER
VOLTAGE STANDARD
SEMICONDUCTOR CONTROL

Fig. 1 (left). Block diagram

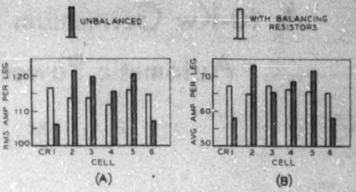


Fig. 3 (right). Rms and average current through rectifying elements with and without balancing circuits

Dual gain-control rheostats *P*7 and *P*8 permit adjustment of the stiffness of regulation of the rectifier. An adjustable stability-control rheostat, *P*13, is provided to compensate for variations in the control loop. Thermistor (*TH*1) minimizes the effects of ambient temperature variations. A feature peculiar to the application in a central office equipped with a pulsed switching load is the combination of capacitor *C*16 and resistor *R*44. This circuit prevents the output voltage from decreasing excessively when the load swings exceed the constant-current settings, for a limited time. If the overload persists, normal constant current protection is realized.

Rectifying Elements

The six germanum power rectifying elements[3] used in this application are rated at 83 amperes per cell in a 3-phase bridge configuration when forced air cooled at the rate of 2,000 linear feet per minute. The quantity of cooling air provided was determined by test to give adequate cooling, while minimizing the noise produced by the high-velocity air passing between the fins of the rectifying cells. Fig. 5 illustrates the blower assembly.

The increased efficiency and small size of the semiconductor power rectifying elements must be accompanied by more complicated protection from short circuits or overloads. Selenium rectifiers have a relatively large surface for the conduction of current and have a time-overload characteristic such that normal fusing provides satisfactory protection. The small area semiconductor cells have a very low thermal capacity, so that the protective device must be extremely fast to prevent damage to the diodes.

The indefiniteness of the term "short circuit" makes the discussion of the subject an involved one. Depending on the impedance of the source, point of short-circuit occurrence, and impedance of the shorting element itself, a short circuit may range from a very slight overload of the

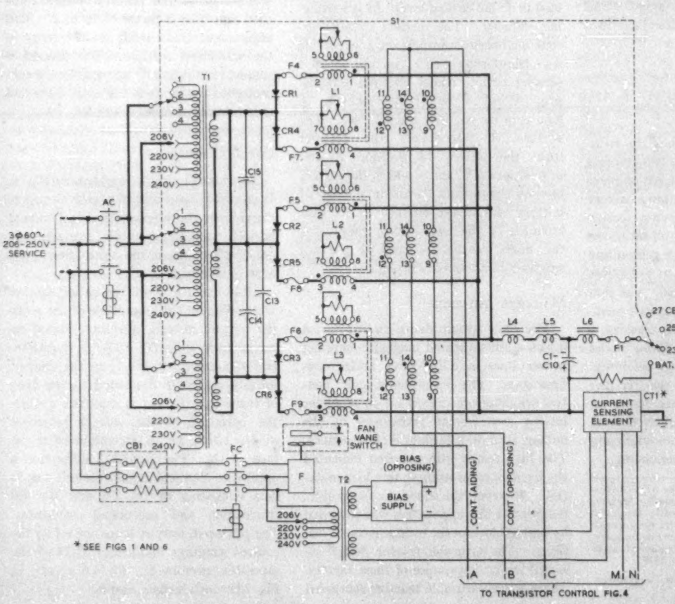

S1
T1
1
2
3
4
206V
220V
230V
240V
AC
3φ60~ 206-250V SERVICE
C15
C13
C14
F4
F7
F5
F8
F6
F9
CR1
CR4
CR2
CR5
CR3
CR6
L1
L2
L3
5
6
7
8
11
12
14
13
10
9
L4
L5
L6
C1-C10
F1
27 CELLS
25 CELLS
23 CELLS
BAT. OFF
CURRENT SENSING ELEMENT
CT1*
FAN VANE SWITCH
CB1
FC
F
T2
BIAS (OPPOSING)
BIAS SUPPLY
+
−
CONT (AIDING)
CONT (OPPOSING)
A
B
C
M
N
TO TRANSISTOR CONTROL FIG.4
* SEE FIGS 1 AND 6

Fig. 2. Schematic of power section

rectifying cell to a current that may be many times the rating of the cell.

In cases where a number of rectifying cells are paralleled and the current capability of the source may be 10,000 amperes or more, the present recommended protection places a surge-type fuse[4] in series with each cell. The currrent-time characteristic of these fuses falls within the safe overload range of the germanium cells. In cases where the short-circuit current is limited by one or more of the factors enumerated above, protection is marginal at best. Protection may then be afforded by a surge-type fuse, a circuit breaker, a current-limiting circuit, or shut-down circuit feature.

In this rectifier three of these methods are utilized to cover specific loading conditions. To protect against short circuits at the rectifying cells, surge-type fuses (*F4-F9* Fig. 2) are used. To protect the cells from overloads external to the rectifier, the electronic control circuit (Figs. 4 and 6) automatically limits the output to a safe value. Failure of the electronic circuit permits an overload relay, *OL* (Fig. 6), to function to shut down the rectifier. However, a short circuit between the rectifying elements and the rectifier output may result in a rectifying cell failure. Since most cells "fail short" a total of three cells can be lost in this circuit. Additional study of this problem is under way since the problem is characteristic of most power supplies whose short-circuit capabilities are low.

Protection against fan failure is provided by a vane-type switch in the air stream. Electrical failure or an obstruction of the intake vent will cause a rectifier shutdown. Voltage surges across the cells are minimized by the use of capacitors across the secondary windings of the main transformer.

CURRENT-SENSING CIRCUIT

The basic current-sensing circuit is the d-c measuring magnetic amplifier circuit[5] illustrated in Fig. 6(A). It has been been called a "d-c transformer" by usage. It is not a new device but its value has been increased by the availability of square-loop core materials.

The device was selected for this application because it provides a signal of considerable magnitude limited only by design. This eliminates the need for a d-c amplifier with its attendant circuit complication and problems of drift, replaceable parts, etc. In this application a 60-volt signal is provided at full load on the rectifier. Fig. 6(B) illustrates the actual circuit arrangement. The lead labeled *L* is the d-c output lead of the

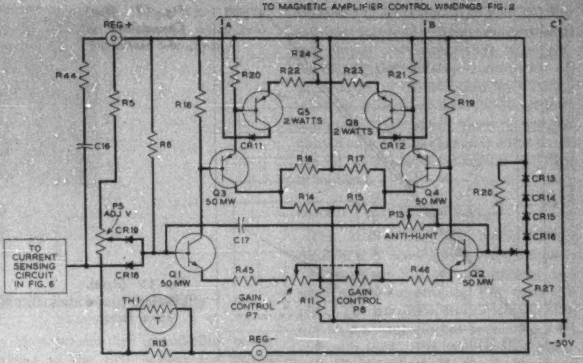

TO MAGNETIC AMPLIFIER CONTROL WINDINGS FIG. 2
A
B
C
REG +
R44
R5
R6
C16
R18
R20
R24
R22
R23
R21
R19
Q5
2 WATTS
Q6
2 WATTS
CR11
CR12
R16
R17
Q3
50 MW
R14
R15
Q4
50 MW
R26
CR13
CR14
CR15
CR16
P5
ADJ V
CR19
CR18
C17
P13
ANTI-HUNT
TO
CURRENT
SENSING
CIRCUIT
IN FIG. 6
Q1
50 MW
R45
R46
Q2
50 MW
R27
GAIN
CONTROL
P7
R11
GAIN
CONTROL
P8
TH1
T
R13
REG-
-50V

Fig. 4. Schematic of control section

Fig. 5. Blower assembly with demountable rectifier assemblies

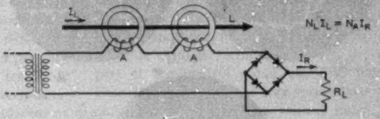

I_L
L
A
A
I_R
R_L
$N_L I_L = N_A I_R$

Fig. 6(A). D-c transformer, basic circuit

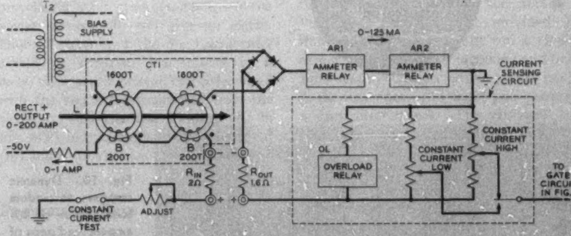

T_2
BIAS
SUPPLY
0-125 MA
AR1
AR2
AMMETER
RELAY
AMMETER
RELAY
CURRENT
SENSING
CIRCUIT
CT1
1600T
A
1600T
A
RECT +
OUTPUT
0-200 AMP
L
-50V
0-1 AMP
B
200T
B
200T
R_{IN}
2Ω
R_{OUT}
1.6Ω
OL
OVERLOAD
RELAY
CONSTANT
CURRENT
LOW
CONSTANT
CURRENT
HIGH
TO
GATE
CIRCUIT
IN FIG. 4
CONSTANT
CURRENT
TEST
ADJUST

Fig. 6(B). D-c transformer, application schematic

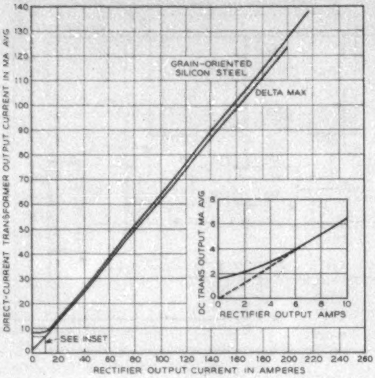


Fig. 7 (left). Characteristic of d-c transformer

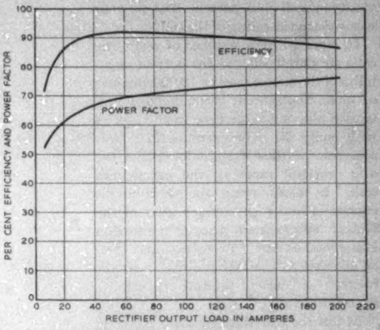


Fig. 11 (right). Efficiency and power factor

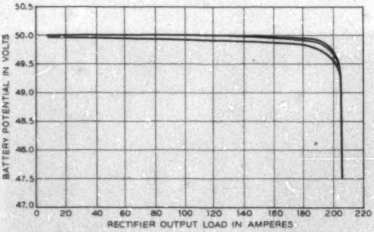


Fig. 8 (left). Static line and load-voltage regulation

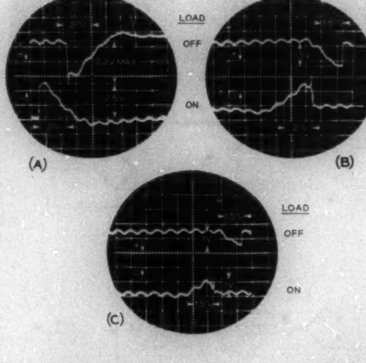


Fig. 9 (left). Dynamic response to a single step-load change

A—95%
B—55%
C—30%

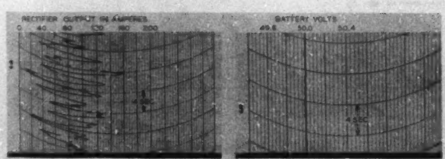


Fig. 10. Dynamic response to random load pulses—six 20-ampere load steps of 0.3–0.5-second duration

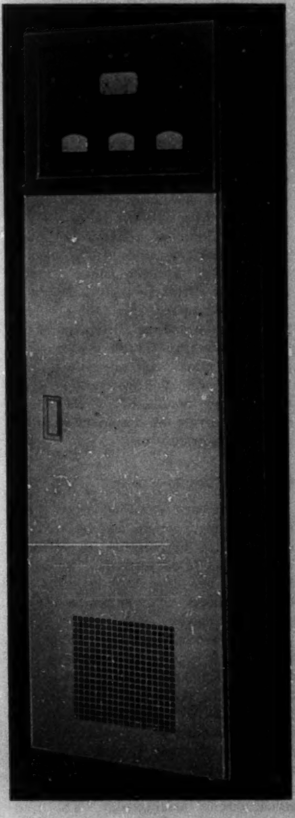


Fig. 12. Model of rectifier

Fig. 13. Plug-in printed wiring control panel

rectifier. It serves as a single-turn primary winding for the toroids making up the current-sensing element *CT*1. The *A* windings are the a-c secondary windings. They are connected series opposing. The ratio of secondary to primary turns is 1,600/1. The rectified output current is proportional to the current in the load lead *L*.

The voltage drop across the group of resistors making up the current-sensing section in the output circuit of the d-c transformer (Fig. 6(B)) is connected to the current gate in the transistor control circuit (Fig. 4). As the load current in the *L* lead increases the voltage across these resistors will increase. If the load current exceeds the preset value permitted by the *CONCUR H* (constant-current high) adjustment, the rectifier output voltage will be decreased to maintain the current at this value. Winding *B* (Fig. 6(B)) is a 200-turn test winding provided to permit adjustment of the constant current level with the rectifier disconnected from the battery and load. A simulated change of 50-150% rated current is possible.

The two saturable cores are Magnetics Incorporated *50042-2A* toroids made of 0.002-inch Deltamax tape. The curve of the rectifier output current, *L*, and the d-c transformer output is plotted on Fig. 7 for a 480-ohm load and ±10% line voltage change. The characteristic for a similar set of coils using grain-oriented silicon steel is also plotted. The principal differences is the low minimum current (magnetizing current) for the high-gain cores.

The saturable-reactors are potted in an epoxy resin for ease of mounting and as physical protection for the windings to avoid abrasion due to contact with the load cable or bus bar. In a similar manner to standard a-c transformer practice, a single design can be used for a series of rectifiers of various current ratings by threading more than one turn through the cores.

Maintenance and Reliability

Most present-day central-office power plants in the Bell System are wholly automatic, and in many cases the offices are unattended. Continuity of service is the paramount objective in the design of telephone systems. To achieve this, a high order of reliability and minimum maintenance of the power plant are essential. The present regulating equipments use electron tubes which are considered renewable items. This rectifier has only components which are considered nonrenewable assuming the components are fabricated reliably. However, to assure prompt restoration of facilities in the event of a failure, a number of indicating devices and trouble shooting aids are incorporated in this design. Three lamps indicate the general area of failure. Circuit breakers or fuses have tell-tale indicators or are self-evident.

The transistor amplifier (Fig. 4) can be tested while the main power section is disconnected from the a-c supply. The transistor amplifier section is on a plug-in printed card making its replacement a simple operation if it becomes defective. The power section can be tested separately by disabling or removing the transistor control section, which also permits manual operation of the rectifier if a replacement transistor-card is not available.

Failure of the bias supply (Fig. 2) will prevent the application of a-c power to the power section and will give an alarm. Faulty behavior of the d-c transformer circuit or apparent inaccuracies in a meter or measuring circuit can be localized with the rectifier disconnected from the battery, using the test winding (*B*) in Fig. 6(B) in conjunction with resistors *R* (IN) and *R* (OUT).

Performance Data

Static and dynamic characteristics of the rectifier model were measured when connected to a 1,000-ampere-hour battery and resistance load. The regulating leads and output voltmeter leads were connected to the battery terminals. All data shown on the attached figures were taken at room ambient (25-30 degrees centigrade). However, measurements indicated only a 0.2% change in output voltage for a 30-degree-centigrade variation in ambient.

OUTPUT VOLTAGE REGULATION

Fig. 8 illustrates the static voltage regulation for a ±7% line-voltage change. The constant current adjustment was made at 205 amperes and 49.0 volts output. The transition from constant voltage to constant current requires a 10% change in load. Fig. 9 illustrates oscilloscope traces of the dynamic response to step changes in load. Response time noted is on a 60-cycle basis. Fig. 10 illustrates the response to a fast changing load simulating operation of the rectifier in a central office with a pulsating relay switching load. The load consisted of six 20-ampere steps appearing at random for approximately 0.3–0.5 second.

POWER FACTOR AND EFFICIENCY

Fig. 11 illustrates the efficiency and apparent power factor of the development model rectifier. Earlier selenium rectifiers used in the same application had a maximum efficiency of 76%.

Equipment Features

Fig. 12 illustrates the appearance of the model rectifier. Note the sloping meter panel to improve readability. The control panels are mounted on a swinging gate behind the large front door. The control amplifier is mounted on a removable printed wiring card which is a part of the control panel, see Fig. 13. Control relays are of the wire spring type, and are mounted on the control panel. The saturable reactors, 3-phase transformer, and filter inductors are mounted in the bottom portion of the rectifier. Ventilated openings are provided in front and rear doors. The top is closed except for openings for incoming and outgoing leads. The air intake for the fan is in the rear door. Air must pass through a dust filter panel. The rectifiers are arranged for side-by-side mounting.

Conclusions

The objective of designing a 10-kw rectifier utilizing solid-state devices, that competes with existing regulated motor-generator sets in telephone power plants, has been realized. Semiconductor devices include high- and low-power germanium rectifiers, low-power silicon diodes, and germanium transistors. The use of laminated core saturable reactors with form-wound coils and an in-circuit balancing arrangement fulfills the requirements of reasonable cost and elimination of core selection while using square-loop core material in a high-gain self-saturated magnetic amplifier. High efficiency and good static and dynamic response resulted from the direct control of the large saturable reactors by a transistor control circuit equipped with two 2-watt transistors in a push-pull output stage. The base area of the unit was decreased approximately 10% over the earlier selenium rectifiers of lower output rating.

References

1. Some Applications of Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers, B. H. Hamilton. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 640–45.

2. The Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, D. H. Smith. *Ibid.*, pp. 645–51.

3. Germanium Rectifiers for Industrial Applications, L. W. Burton. *Ibid.*, vol. 75, pt. I, Mar. 1956, pp. 41–44.

4. Shawmut Amp-Trap. *Bulletin 514-7*, The Chase Shawmut Company, Newburyport, Mass.

5. Magnetic Amplifiers (book), H. F. Storm. John Wiley & Sons, Inc., New York, N. Y., 1955.

Silicon Power Rectifier Equipments

J. A. MARSHALL
MEMBER AIEE

V. N. STEWART
MEMBER AIEE

A NOTABLE newcomer in the power rectifier field is the silicon power rectifier cell capable of handling, in some applications, power outputs up to 10,000 watts each. This paper illustrates one method of using silicon power rectifier cells in an equipment with an industrial rating of 300 kw at 250 volts and with overload duty of 125% for 2 hours and 200% for 1 minute. A review of the components, equipment design, and operating performance is presented along with general rules and their application to equipment protection.

Circuit

The double-way or 3-phase full-wave bridge rectifier, Fig. 1, proved to be the most economical for this application because it allowed the use of simplified copper configurations and a standard transformer.

Components

The silicon power rectifier cell was chosen for its peak inverse- and surge-voltage characteristics and higher operating temperature. The type *4JA60B* cell[1] most closely met the requirements of the normal peak inverse voltage of 276 volts and the surge duty of 600 volts as established in the section on application basis for components. The magnitude of the surge voltage demanded two cells in series, and current requirements dictated the total number of cells per phase. There are a total of 192 cells used with 16 parallel rectifying paths for each way, and two cells in series for each parallel path. The average current per cell with this arrangement is 25 amperes for the full-load rating of the equipment.

A-c power is supplied to the rectifier unit by a standard 500-kva delta-Y 480-volt to 208/120-volt transformer with four 2% taps. Voltage regulation is accomplished by an induction-type regulator rated 36 kva, 480 volts, capable of controlling the output voltage ±10%.

For best design, the cells should be mounted with consideration given to simple circuit connections, adequate cooling, and ease of maintenance. A removable tray was designed to accomplish the foregoing goals. Eight silicon power rectifier cells, their dividing resistors, a protective fuse, air baffles, disconnects, fin conductors, an indicating lamp, lift-off handle and latch are arranged on a tray as shown in Fig. 2.

The eight cells of each tray are connected electrically as shown in Fig. 1 and physically as shown in Fig. 2. The incoming and outgoing connections are terminated in finger-type disconnects.

The five copper fin conductors to which the cells are connected act as heat sinks and the area between them allows for the passage of cooling air, either natural or forced convection. Four of the fins provide the common point between series cells and the fifth fin is the common conductor for the four parallel paths.

The indicating lamp, connected in parallel with the fuse, lights whenever the fuse blows and the circuit is energized. It is possible to include an audible alarm or remote indication of fuse operation for unattended installations.

Three vertical rows of four trays each are mounted on the front and rear of the metal-enclosed unit which houses the bus. The bus consists of the three a-c power connections from the transformer and the two d-c outputs to the load air circuit breaker.

The unit shown in Fig. 3 has a fan mounted on the roof to provide forced air cooling for the trays.

Normal maintenance of the trays consists of a visual inspection and a check of the internal connections. This is quickly accomplished by disconnecting the tray as shown in Fig. 4. In the event of a cell failure, quick replacement of the tray having a faulty cell is possible under load. A spare tray can be inserted while necessary replacements are made to the original tray.

Protection

A completely protected rectifier equipment is accomplished through a design combining the characteristics of the rectifier cells, transformers, fuses, and breakers. These components form an equipment to give a continuous-current rating and short-time ability as well as providing for the following abnormal conditions of an industrial system:

1. D-c overload.
2. External d-c faults.
3. Internal d-c faults.
4. Cell-blocking-characteristic failure.
5. A-c faults.

Applying protection under these conditions has established several general rules which give high continuity of service, rapid restoration of service after an overload outage, short-circuit protection, and removal of a defective cell. This applies particularly to power units having

---

Paper 57-1176, recommended by the AIEE Industrial Power Rectifiers and Semiconductor Metallic Rectifiers Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference on Rectifiers in Industry, Chicago, Ill., June 4–6, 1957. Manuscript made available for printing September 20, 1957.

J. A. Marshall and V. N. Stewart are with General Electric Company, Philadelphia, Pa.

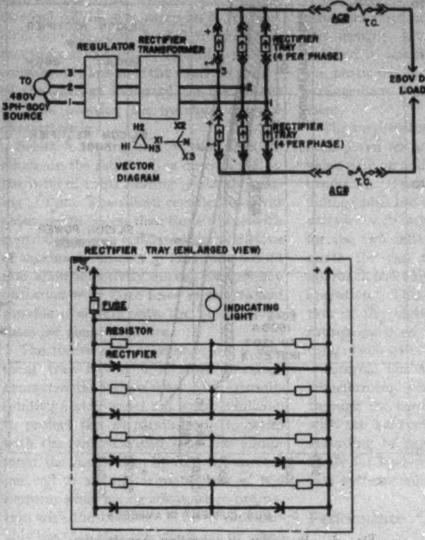


Fig. 1. Double-way silicon power rectifier circuit

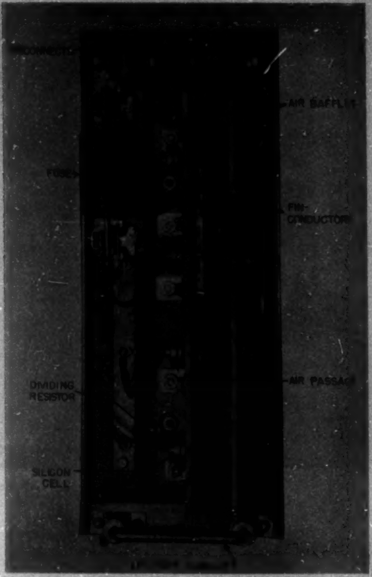


Fig. 2. Removable tray showing type SRF rectifier fuse and type 4JA60 silicon power rectifier

multiple rectifier cells per phase and capacities such as the 300-kw unit presented in this paper.

The following general rules can be partially demonstrated in the co-ordination curves of Fig. 5 and in the basis for selection of the components which follow this section.

1. Low thermal capacity of the rectifier cell makes the overload duty an influencing factor in the determination of unit size.

2. Momentary capacity of the equipment (rectifier cells) must be sufficient to withstand fault durations equal to the clearing time of the d-c line breaker.

3. Maximum fault conditions cannot exceed the surge-current limits of the cells without a probable loss of cells.

4. Fuse melting characteristics must allow sufficient time for circuit clearing by the d-c breaker over the entire range.

5. A fuse in series with a rectifier cell or parallel cells as used in this equipment can only provide cell isolation with a failure of the blocking characteristic. In general, fuses with a sufficient current capacity to meet the normal forward duty will not provide individual cell protection on overload or short-circuit currents. Further, with several fuses operating in parallel under balanced conditions and appearing to divide the available fault current, they will not clear simultaneously and thereby leave the total duty to the last fuse to clear. This means that the maximum duty of the cells must be based on the entire fault through the last fuse to clear. The extreme duty on the cells protected by the last fuse to clear would result in thermal damage and loss of the cells.

6. Fuse selectivity under failure of the blocking characteristic is dependent on the fuse-clearing, arcing, and melting i^2t characteristics established for maximum system conditions. Performance for this function cannot be evaluated from the normal time-current melting characteristic of the fuse but must include the energy conditions with full voltage applied.

7. Surge voltages occurring during the operation of protective equipment as well as the normal voltage conditions will determine the selection of the operating peak inverse voltage rating of the rectifier cells.

Application Basis for Components

The rectifier equipment presented in this paper readily illustrates the selection of components based on their characteristics and the requirements established for adequate protection. Using the system shown in Fig. 1, an equipment supplying a 300-kw load at 250 volts will require 16 parallel conduction cells in each phase to meet both the industrial overload and short-circuit momentary ability.

Fig. 3. Front view of the silicon rectifier equipment with type AK-1-50 air circuit breaker

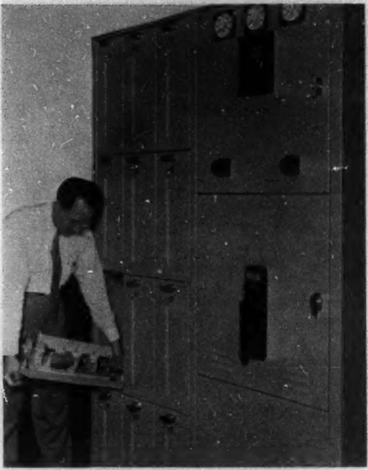


Fig. 4. Method of tray removal

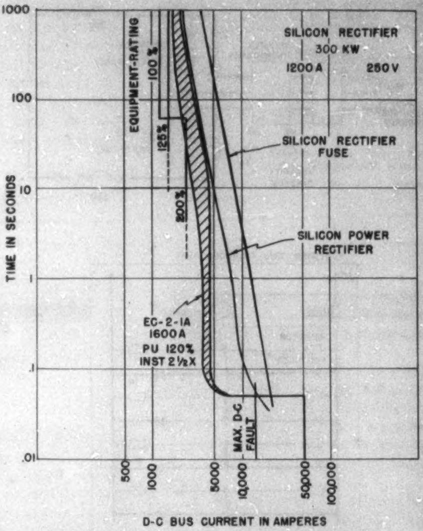


Fig. 5. Protective co-ordination characteristics

Each cell is applied to limits determined by cell heating due to losses, heat-dissipation rate, ambient conditions, thermal storage capacity of conducting parts, and thermal capacity of the cells.

In the selection of the number of cells a conservative continuous-current level is established when the numbers of cells are adequate for the overload protection associated with the 125% and 200% ratings of the industrial equipment. The 16 parallel cells of this unit have the characteristics of the silicon power rectifier in Fig. 5 and the long-time overload duty to the cells which is determined by the protective breaker characteristic can reach a maximum of 1,980 amperes (165%). For these overloads with very long times, the output of the cells is dependent on the power loss and the ability of the cells and equipment to dissipate this loss. Although the initial design data will be based on a desired copper conductor temperature for calculating the selected arrangement, such as illustrated in Figs. 1 and 2, it is necessary to evaluate the results through a test of a prototype and the completed equipment. Each degree rise in the bus limits the average current capacity of the cell which has a total operating temperature of 200 C (degrees centigrade). The cell temperature of this equipment is based on 48 cells averaging the normal d-c output and is related to the power losses of the 3-phase curves as given in a previous paper.[1]

As the values of overcurrent increase, the time to reach the temperature of the cells will reduce. Here, the rectifier characteristic has a resemblance to that of a fuse characteristic with an inverse time function in the range of 10 to 600 seconds. This damage characteristic for the rectifier cell must include the initial temperature conditions at 100% load, the maximum ambient of 40 C, the heat sink provided by both cooling fins and the rectifier parts, and the temperature generated by the junction losses.

As the peak currents increase to short-circuit conditions, the damage characteristic shifts from a temperature limitation to one of cell fracture under high currents. This alters the characteristic as can be noted for values less than 1/2 second. This has been verified through surge testing of individual and multiple cells as well as power tests of the entire equipment.

For this application the available faults levels are from 9,000 to 11,000 amperes d-c which is well below the momentary ability of the cells with a breaker clearing time of 0.05 second. The equipment with the 16 parallel cells per phase has momentary ability approaching 15,000 amperes.

The melting time-current characteristics of the fuse must be co-ordinated with the d-c breaker over the entire range of currents. This co-ordination is evident in the protective co-ordination curves for this application for values over the range and beyond the maximum d-c fault conditions.

The fuse conduction period in each cycle is related to the commutation characteristics of the cell which change from 120 electrical degrees conduction for the lower overload current to 180 electrical degrees for the fault conditions. Therefore, the normal fuse characteristic must be properly adjusted to show its current melting response in relation to the rectifier applications, and this has been made in Fig. 5.

Blocking characteristic failures cause direct line-to-line faults on the transformer through the normal conducting cells of another phase and must be isolated by limiting action to prevent damage of the good cells. This is the function of the fuses applied with the equipment and the minimum number of fuses, such as the four per phase used in this 300-kw unit, are determined by the contingency operation with one fuse open, the relation of total clearing energy to melting energy, and the peak let-through currents. The last two conditions are vitally affected by the system's voltage and available currents. For the contingency operation with one fuse open, a phase with a defective cell will have to carry an increase of 33% over the 100% load during replacement of cell and fuse. This duty will be considerably less than the overload ability of the cells as indicated by the earlier analysis covering the overload damage characteristic.

Selective operation by the fuse to eliminate the failed cell is determined by the ratio of total clearing $\int i^2dt$ to melting $\int i^2dt$. The silicon rectifier fuse was measured to be less than three to one with controlled power testing using conditions of maximum severity. This ratio would also allow selectivity during contingency operation with three fuses in service and provide a margin with the normal four fuses per phase in service.

The fuse which is isolating an equipment tray having a rectifier blocking-characteristic failure must give current-limiting action under the worst condition to protect the 16 parallel cells in series with the faulted circuit. In this equipment the fuse design limited the current per cell to an approximate peak of 650 amperes while giving adequate co-ordination with the breaker. This is considerably less than the value which can be taken by the cells for several cycles. In this application the $\int i^2dt$ let-through energy is one fifth of the ability of the silicon cells and does not represent a factor in the design; however, this should be checked to assure sufficient margin with the particular rectifier characteristic and arrangement being used for the equipment.

The surge-voltage rating of the cells has been given as nominal continuous peak value with an additional amount for surge condition. The 200 volts of peak inverse voltage plus 100 volts for surge makes 300 volts for each unit and a total of 600 volts for the two cells in series. This is adequate for the maximum peaks which approach 600 volts for the fuse or breaker operation. This eliminates the consideration of the higher peak inverse voltage ratings such as the 300-volt cell with a total of 400 volts for surge conditions.

Internal faults in the equipment and transformers receive their protection through the breakers or fuses associated with the a-c system. These are applied according to practices now used extensively for load-center units and distribution systems supplying a-c power.

Performance

An induction-voltage regulator with a range of ±10% is connected ahead of the main transformer and provides voltage regulation. The inherent regulation of the unit without the induction regulator is approximately 5%.

Using a forward-cell drop of approximately 1 volt, the efficiency of the rectifier element is approximately 98% and the efficiency of the equipment per American Standards Association mercury-arc rectifier standards is approximately 95.5%. The over-all efficiency including the induction-voltage regulator required for voltage control is approximately 94.3%. The vector power factor is approximately 97%.

Conclusions

The performance of this equipment compares most favorably with other types of conversion equipment in the 250- to 500-volt range. It is felt that the voltage range will be extended higher in the near future. This type of equipment is best suited for use in industrial and electrochemical applications which do not require large voltage ranges.

Reference

1. THE RATING AND APPLICATION OF A SILICON POWER RECTIFIER, D. K. Bisson. *AIEE Special Publication T-93*, "Rectifiers in Industry," June 1957, pp. 135–42.

Selenium Rectifier Applications in Automotive Vehicles

EMERY J. SZABO
ASSOCIATE MEMBER AIEE

SELENIUM rectifiers were first tested in 1937 for possible application in automotive vehicles. The term "automotive vehicle" includes passenger cars, trucks, buses, off-highway construction equipment, and other self-propelled vehicles.

It was in 1937 that the Chrysler Corporation was doing experimental work with an alternator rectifier system, replacing the conventional d-c generator on passenger cars. The first selenium rectifiers for that application were imported from Europe since they were not manufactured in the United States until 1938. There was considerable development work done by the Chrysler Corporation on passenger car systems but the beginning of World War II stopped the commercial work and Chrysler continued development of an alternator rectifier system for military aircraft.[1] There was limited production of an aircraft alternator rectifier system rated at 30 volts, 400 amperes d-c output. The selenium rectifiers were air-blast-cooled for minimum size and weight.

In 1945, the Leece Neville Company took up development work on an alternator rectifier system for heavy duty applications on police cars, trucks, and buses.[2]

The first standard equipment application of selenium rectifiers was in 1946, when the Twin Coach Company introduced their postwar design motor coach, which was equipped with an alternator rectifier system rated at 14 volts, 100 amperes d-c output. Since that date, the use of selenium rectifiers in vehicle generating systems has expanded rapidly. Up until today there have been about 100,000 selenium rectifiers installed as a part of heavy-duty generating systems for automotive vehicles.

Before discussing the particular applications of various types of selenium rectifiers in generating systems, a comparison will be made of the conventional d-c generating system with the alternator rectifier system. The system consists of two units; a variable speed-belt-driven d-c generator, and the generator control unit, commonly referred to as the voltage regulator. The d-c generator is a shunt-type machine and is required to operate over a speed range established by the pulley diameters and speed range of the engine in the vehicle.

The alternator rectifier system consists of three units. The first of these is the

---

Paper 57-1177, recommended by the AIEE Industrial Power Rectifiers and Semiconductor Metallic Rectifiers Committees and approved by the AIEE Technical Operations Department for presentation at the AIEE Conference on Rectifiers in Industry, Chicago, Ill., June 4–5, 1957. Manuscript made available for printing September 20, 1957.

EMERY J. SZABO is with the Leece Neville Company, Cleveland, Ohio.

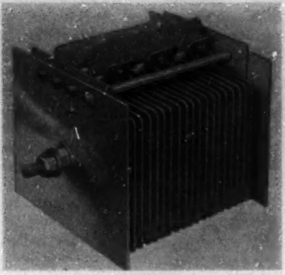


Fig. 1. Fourteen-volt 100-ampere selenium rectifier, 1946

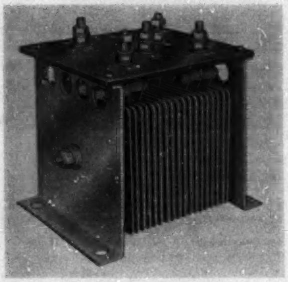


Fig. 2. Heavy-duty railroad type of selenium rectifier, 14 volts, 125 amperes

belt-driven variable speed alternator which operates over a speed range established by the pulley diameters and the speed range of the engine. The second unit is the rectifier which converts the a-c output of the alternator into d-c power for battery charging and electrical loads on the vehicle. The third unit is the generator control unit which is very similar to the d-c generator control unit. The major difference between the d-c generator control unit and the alternator rectifier control unit is in the cutout element. The control unit for the alternator rectifier system does not have a reverse current winding on the relay, since there is no appreciable amount of reverse current discharged from the battery through the rectifier. In most applications the relay is energized or de-energized from the ignition switch on the vehicle so that the relay remains closed at all times during operation of the vehicle.

Basically the two systems serve the same function of providing electric energy to charge the battery and carry the electric loads on the vehicle. A comparison of the output characteristics of the two systems, however, shows considerable difference in the power which may be obtained from the alternator rectifier system compared to the output of the d-c generating system. There are two distinct advantages for the alternator rectifier system. One of these is more power available at lower speeds for minimum size and weight of the generating unit and other is the ability of the system to operate over a wide speed range. The alternator itself is about half the weight of a d-c generator required to generate the same amount of power at the same minimum full-load speed of the unit. The other major advantage of the alternator rectifier system is the ability to produce relatively large output at low speed by taking advantage of higher belt drive ratios, and at the same time its ability to generate this high output at maximum engine operating speed with good reliability even at maximum load and speed. This permits design and application of alternator rectifier systems to produce up to 50% of maximum rated output at engine-idling speed and still be able to operate at maximum speed and output without problems of commutation and short brush life. The problems of commutation at high speed require special design considerations on d-c generators which are usually prohibitive from the cost standpoint in automotive vehicles.

The selenium rectifier is the device that permits the design of the alternator to be independent of the wide speed range and high drive ratios which are required in today's vehicle generating systems.

The selenium rectifier, being a separate unit, is located independently of the alternator and is usually mounted either directly behind the engine radiator cooling fan or in front of the engine radiator. In both of these locations, the cooling air is provided at all times while the engine is running. The rectifier is used at maximum air-blast-cooled ratings in order to keep its size, weight, and cost to a minimum. There have been a few applications of convection-cooled rectifiers such as on railraod caboose cars, but the size and cost of the convection-cooled rectifier usually prohibits its use where air-blast cooling is available. Another method of cooling the rectifier is by locating it within a duct work which is either in the engine air intake system on diesel engines or which leads to a rectifier cooling fan on the alternator. This fan draws air through the duct work and thereby provides air-blast cooling for the rectifier whenever the alternator is rotating. The engine air-intake cooling method mentioned previously has proved very successful on large diesel engines used in tractor-trailer combinations and also in off-highway construction equipment such as 30-ton dump trucks.

Fig. 3. Alternator fan duct-cooling of selenium rectifier

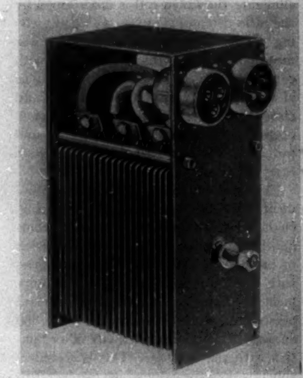


Fig. 4. Military type of selenium rectifier rated 28 volts, 100 amperes output

The application of selenium rectifiers in automotive vehicles presents two problems which do not usually exist in industrial applications. One of these is the problem of mechanical vibration, since the rectifiers are mounted where they are subjected to considerable vibration from the engine as well as road shock. This has required special mechanical design features in the selenium rectifiers. There has been the necessity for using a keyed construction for the plates and terminals on the center stud to keep them from rotating under vibration and causing a failure. Also there has been a special support of the terminals themselves to keep them from failing under vibration. The other major problem encountered by the selenium rectifier in automotive vehicle applications is that of exposure to extreme corrosion-causing elements. Most of the selenium rectifiers are mounted in such a location that they are exposed to dirt, road splash, and worst of all, the tons of salt spread on the highways during the winter. The road salt mixed

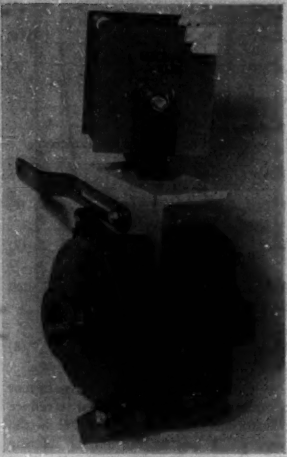


Fig. 5. Selenium rectifier mounted on top of alternator

with the slush on the pavement has required special salt spray resistant finishes on the selenium rectifiers. It is standard practice to specify the salt spray resistant finish on all selenium rectifiers used in automotive vehicle applications and even then there are still occasional failures from corrosion. The few applications which might be adequate with ordinary paint finishes do not justify the complications involved in deciding what type of finish would be good enough for various applications. These two problems have been successfully solved so that today the percentage failures are low enough to be of no concern.

The 14-volt 100-ampere selenium rectifier which was the first production application is shown in Fig. 1. This is a 3-phase bridge stack consisting of 18 cells $4^1/_8$ by 6 inches. The terminals, in addition to being keyed on the center stud, are also supported by insulator tubes to prevent failures from vibration. About 30,000 of this particular type of rectifier have been placed into service. The cells, although operating at $3^1/_2$ times normal convection rating, have given very satisfactory service.

The rectifier shown in Fig. 2 was introduced for railroad and marine use, as indicated by the rugged mechanical construction features. There are two heavy steel end plates with an insulator panel across the top supporting the large nickel-plated studs to which the leads are connected. This rectifier is rated at 14 volt, 125-ampere output, and uses 18 cells of 5- by 6-inch size. Fig. 3 shows the duct alternator fan cooling method for a rectifier of the construction shown in Fig. 2.

The next rectifier, Fig. 4, is a military unit rated at 28 volts, 100-ampere output. It consists of 18 cells, $4^1/_8$ by 6 inches, connected in a 3-phase bridge. Thousands of these are in service as part of heavy-duty alternator rectifier generating systems on military vehicles.

The combination of units, Fig. 5, shows a rectifier mounted on top of an alternator with adjustable rectifier-bracketing to permit positioning the rectifier for maximum cooling air from the radiator cooling fan.

A new type of selenium rectifier used in automotive vehicles is the one shown in Fig. 6. This uses the high current density selenium cell which permits further reduction in size and cost for the rectifier in automotive vehicle generating systems. This rectifier was introduced in 1954 with a rating of 14 volts, 50 amperes' output. It consists of six high-current-density cells of 4 by 4 inches. Comparison with previous units shown establishes the major step forward in rectifier design for smaller size and lower cost. The availability of this rectifier has been a major factor in producing a lower cost alternator rectifier system which has proved so popular for automotive applications that already about 50,000 of these rectifiers have been put into service.

The high-current-density cell having proved successful in the 14-volt 50-ampere rating is now being applied to the 14-volt 100-ampere and 150-ampere ratings in order to provide lower cost units to replace those described previously. For example, the use of the high current density cell will permit design of a 14-volt 200-ampere output rectifier no larger than the 14-volt 100-ampere heavy duty unit shown previously in Fig. 2. The high-current-density cell is quickly replacing the conventional cells for all automotive vehicle applications because of its higher output and lower cost, the two major requirements in these applications.

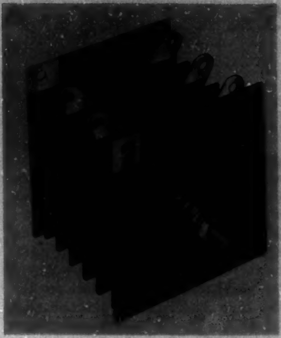


Fig. 6. High-current-density 14-volt 50-ampere selenium rectifier (1954)

The alternator rectifier system has a potential future of becoming standard equipment on all passenger cars even though its initial cost is somewhat higher than that of the conventional d-c generating system. The requirements for higher outputs, along with the more important requirement for output at engine-idling speed as well as reliability at high speed will promote the use of alternator systems within the next few years as standard equipment. The selenium rectifier plays an important part in the applications of these systems because in the past the lack of a high-output, small-size, low-cost rectifier was the main problem in applying alternator rectifier systems to vehicle usage.

References

1. SELENIUM RECTIFIERS IN MOTOR VEHICLE POWER SYSTEMS, Glen Ramsey. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 149–53.

2. AN A-C VEHICLE GENERATING SYSTEM, Albert D. Gilchrist. *Electrical Engineering*, vol. 66, Aug. 1947, pp. 779–80.

Size Reduction of Air-Borne Transformers

RAY E. LEE
NONMEMBER AIEE

SIZE and weight are very important design factors for air-borne electronic equipment. The cost of adding a pound of weight to the electronic equipment is $400 per airplane or more.[1] One very effective way to reduce electronic equipment weight is to reduce the weight of the power transformers. Much progress has been made in this direction since World War II. In 1950 17 pounds were saved in one radar set by redesigning the transformers to use improved oriented steel cores. This represented a saving of 40% of the transformer weight of this equipment. Greater savings of space and weight than this are still possible through improved transformer design.

In this paper the methods for transformer size reduction are enumerated and analyzed. Sample designs which illustrate some of the most promising techniques are described.

Factors Affecting Size

Several of the methods for reducing transformer size and weight seem obvious, but are mentioned here because they are often neglected. For instance, 2.7 pounds were recently cut from one transformer simply by carefully re-evaluating the load and then redesigning to a more realistic value. Simplification of mounting, simplification of terminals, better physical tolerances, more exact knowledge of design limitations are all means that can result in very valuable weight and size savings. The care with which a design is made can often reduce its size as much as 25%. Large tolerances are often excused on the basis of cost, but if total cost of the increased weight were considered the cost of the closer tolerance would be justified.

Often possibilities for weight reduction are overlooked because the equipment engineer does not realize how his design affects transformer size. For instance, a plate transformer for d-c power supply will be 17 to 25% smaller if a bridge-type rectifier is used in place of one requiring a center-tapped transformer. Power supplies using capacity input filters or half-wave rectifiers will in general require larger transformers than are needed with power supplies employing choke input filters and full-wave rectification. Solid-state rectifiers and 3-phase systems result in reduced magnetic component size, because the filament transformers are eliminated, the filter chokes can be made much smaller, and the 3-phase transformer is smaller than its single-phase counterpart. Thus, considerable weight and size reduction is often possible without considering any basically new techniques.

THE GENERAL TRANSFORMER EQUATION

The general transformer equation, discussed by Garbarino, will be used to show how various parameters affect transformer size.[2]

$$W_r = 1/45 f F_c F_i B \Delta A_c A_i \qquad (1)$$

where

A_c = area of core window in square inches
A_i = core cross-sectional area in square inches
W_r = load volt-amperes of transformer
f = frequency in cycles per second
F_c = winding space factor, the fraction of total window area occupied by conductor cross section
F_i = core space factor
B = maximum flux density in kilolines per square inch
Δ = current density in kiloamperes per square inch of conductor

By using the best ratio of copper to iron and by using the best configuration for the core and coil some size and weight reduction can usually be accomplished. Garbarino discusses this in detail in his report.[2]

If the configuration of the transformer is held constant while other parameters are varied, then the affect of these parameters on transformer volume can be investigated. When all other size-reduction methods have been considered, an optimum configuration can then be computed to give either minimum volume or minimum weight.

To obtain an expression for volume, equation 1 is rewritten:

$$A_c A_i = \frac{45 W_r}{f F_c F_i B \Delta} \qquad (2)$$

$$V \propto l^3 \qquad (3)$$

$$A_c \propto l^2 \qquad (4)$$

$$A_i \propto l^2 \qquad (5)$$

where

V = volume
l = any linear dimension of the transformer

Then

$$V \propto (A_c A_i)^{3/4} \qquad (6)$$

$$V \propto \left(\frac{45 W_r}{f F_c F_i B \Delta}\right)^{3/4} \qquad (7)$$

A configuration coefficient is defined such that:

$$V = C\left(\frac{W_r}{f F_c F_i B \Delta}\right)^{3/4} \qquad (8)$$

Equation 8 shows which transformer parameters can be manipulated to reduce transformer size, but it does not show the penalties in loss, regulation, or heating that are incurred. These are discussed later. The change of size with configuration will not be discussed here since it is discussed adequately elsewhere.[2]

By examining each of the factors at the right side of this equation, it is easy to see what is necessary to reduce the size of a transformer.

Of course, if W_r, transformer output, is reduced the size is reduced. Some of the factors in the denominator of equation 8 must be increased to reduce the size of the transformer further.

If frequency is increased the size of the transformer can be greatly decreased, but the transformer design engineer rarely can control this factor.

Winding space factor usually varies from 0.10 to 0.40. There are great possibilities of decreasing the transformer size by increasing the winding space factor. This can be accomplished with improved manufacturing methods and improved insulation. Foil-wound coils show considerable promise as a technique for improving coil space factor.

Size can be reduced only slightly by increasing core space factor since it usually is already 0.9 to 0.95, and the maximum factor is 1.0.

Transformers are usually designed so that the flux density is a little less than that required to saturate the core; therefore, B cannot be increased until a better core material is made available. Cobalt iron alloys show promise of increasing the permissible operating flux density by 25%.

Current density can be increased until loss, regulation, or heating becomes a limiting factor.

Paper 57-1013, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 7–11, 1957. Manuscript submitted July 15, 1957; made available for printing August 23, 1957.

RAY E. LEE is with the Westinghouse Electric Corporation, Baltimore, Md.

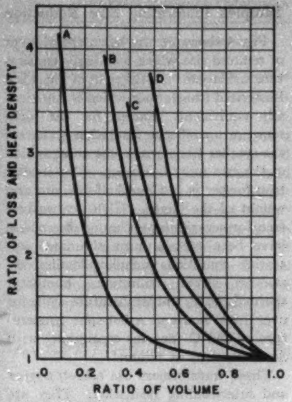


Fig. 1. Ratio of loss and heat density increase as transformer size is reduced

A—Loss ratio as frequency is increased
B—Heat density ratio as frequency is increased
C—Loss ratio as current density is increased
D—Heat density ratio as current density is increased

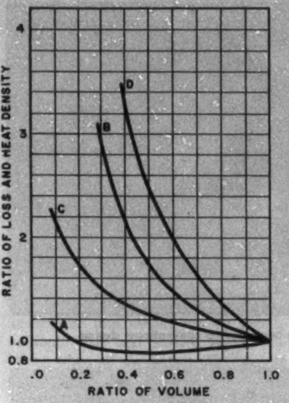


Fig. 2. Ratio of loss and heat density as transformer size is reduced

A—Loss ratio as coil space factor is increased
B—Heat density ratio as coil space factor is increased
C—Loss ratio as current density and flux density are increased at the same rate
D—Heat density ratio as current density and flux density are increased

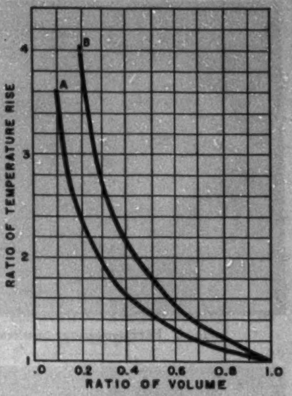


Fig. 3. Ratio of temperature rise as transformer size is reduced

A—Temperature rise when total heat flow is constant
B—Temperature rise when transformer is reduced by increasing current density and flux density at the same rate

Loss and Heating

It has been shown that smaller transformers can be built if the frequency, winding space factor, flux density, current density, or any combination of these are increased. To reduce the size will usually increase the loss, but it will increase the heating even more since the increased heat is dissipated from a smaller volume. Figs. 1 and 2 show how the loss and heat density are affected as size is reduced by increasing various parameters. These curves give the ratio of loss and heat density in the transformer of reduced size to the loss or heat density of a standard design of the same configuration. It is assumed that in the standard design the loss in the coil is equal to the loss in the core; no other limitations are necessary on the standard design. Equations for curves are derived in the appendix.

The watts per cubic inch dissipated as heat are designated as heat density. The temperature rise for any configuration will increase as the heat density increases, but it will not be directly proportional to the heat density. The heat distribution between core and coil will usually change as various parameters are changed. If the heat distribution remains the same within the transformer, then the temperature rise for a given heat density will be less with a small transformer than with a large one. Fig. 3 shows how temperature rise increases as the transformer size is reduced. Curve *A* assumes the total loss remains constant and curve *B* is for the case where size reduction is accomplished by increasing B and Δ in the same ratio.

A few examples may show more clearly how loss and heating are affected by size reduction. If a transformer design is reduced to 70% of its former volume by increasing the current density, then it will have 50% more loss. The heat density will increase 79%, but most of the loss will be in the coil. If the first design was 96% efficient, then the new design will be 94% efficient. If this same reduction is accomplished by increasing both B and Δ in the same ratio, then the loss is increased only 13%; the heat density is increased 61% and the temperature rise is increased 36%. If 30% reduction is accomplished by increasing current density and winding space factor together in equal amounts, then the loss would increase 2%, heat density 46%, and temperature rise 12%.

Heat Removal

To decrease the size of a transformer will in general increase the heat density within the transformer and this will raise the temperature unless better cooling is provided. The heat density is increased primarily because the loss is dissipated from a smaller volume. The small decrease in transformer efficiency will not usually be a limiting factor in the design.

All the heat due to loss must be removed from the transformer. There are in general two ways to dispose of the same or more heat from a smaller volume. One method is to use high-temperature insulation and let the transformer run hot. If the channels for heat flow are not improved, the temperature rises until heat flow due to the temperature gradient equals the heat originating in the transformer. It should be remembered that the higher temperature is not primarily the result of more heat, but the same heat in a smaller space; therefore, the ultimate heat sink is not increased significantly. The problem is getting the heat to that heat sink. The other method of heat removal is to provide for a better heat-conducting path between the heat source in the transformer and the ultimate heat sink. Perhaps the best design for many applications is a combination of these two heat-removal methods. Thermal conductors built into the transformer can greatly improve the heat flow and, thus, reduce the temperature rise in the transformer. Circulating or evaporating fluids are other very effective ways to obtain increased heat flow from the transformer.

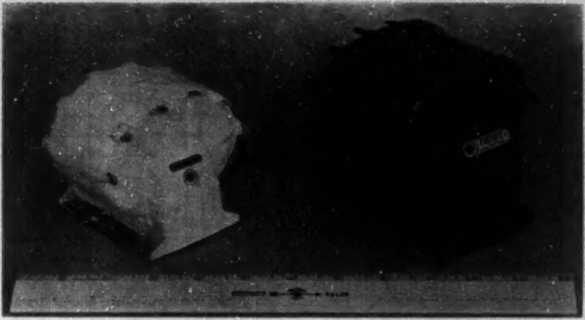


Fig. 4. High-temperature transformer compared to conventional design

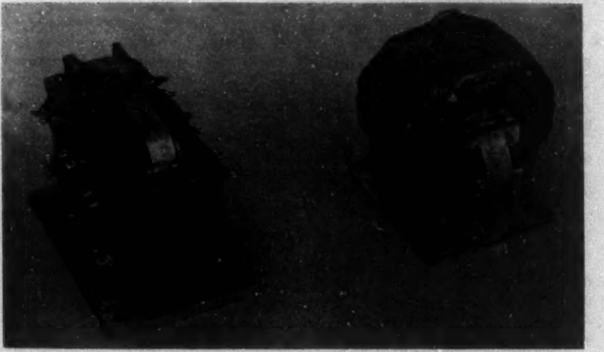


Fig. 5. Three-phase transformer using thermal conductors compared to one phase of conventional 3-phase design

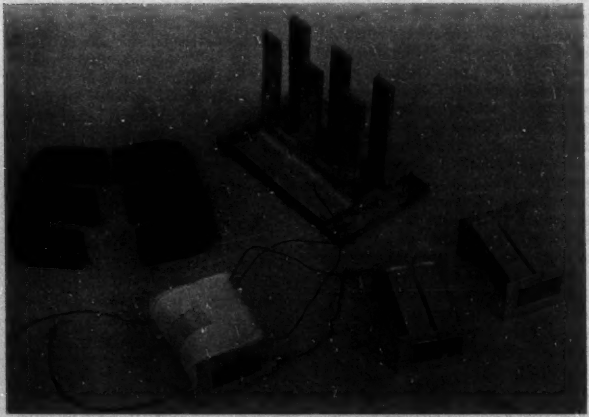


Fig. 6. Construction of transformer with thermal conductors

Samples Illustrating Size Reduction

Fig. 4 shows on the left a transformer of reduced size which uses high-temperature insulation. At the right of Fig. 4 is a Fosterited transformer which is functionally equivalent. Except for the high-temperature silicone insulation system of the smaller transformer, the construction of the two units is similar. The smaller transformer is but one half the size and weight of the larger. High-temperature transformers similar to the one shown here have been tested under simulated operating conditions including loading, dielectric stress, vibration, humidity, thermal shock, and temperature aging for over 1,000 hours at a hot-spot temperature of 250 C (degrees centigrade) with no indications of failure.

These transformers use asbestos layer and interwinding insulation. They are encapsulated with Silastic rubber and vacuum-impregnated with a solventless silicone resin. The solventless silicone insures a thorough impregnation of the coil, while the Silastic provides for a high moisture resistance on the outside of the transformer and excellent insulation between terminals.

Fig. 5 shows on the left a 3-phase transformer of reduced size which uses metal thermal conductors to improve heat transfer in the transformer. The transformer on the right of Fig. 5 is one of three single-phase conventional transformers needed to do the same job. The 3-phase unit is built so that thermal conductors in the transformer carry the heat to the base and from there the heat passes into a "cold plate."

Fig. 6 shows the construction of the transformer. The frame is made of aluminum with the vertical members designed to conduct heat from the coil and core to the base. The coil is wound on an aluminum form with a slot to prevent a short-circuited turn. The flanges at the ends of the form keep the form from collapsing and also help to retain the wire. These flanges are insulated with split insulating washers that fit over the coil form. The core is fitted between the vertical members of the frame and the coils are fitted onto it. Close thermal contact is kept between the core, the coil, and these vertical members. Electrical insulation is insured by a tough anodized surface. The heat originating in the core flows readily through the iron of the core into the frame conductors and from there to the base. Likewise, heat from the coil flows through the thin coil section to the coil form and from there to the vertical conductors and then to the base. The base of the transformer must

Table I. Three-Phase Transformer Performance

Total transformer output = 590 volt-amperes
Total transformer loss = 49 watts
Transformer efficiency = 92.5%
D.C. output with silicon rectifiers = 320 volts at 1.65 amperes = 528 watts

be mounted on a cold plate. About 4 watts per square inch must be removed from the transformer base. The cold plate need not be very cold. In this instance 125 C was used and served its purpose well. The temperature rise above the base in such a transformer can be computed by adding all the incremental temperature rises along any path from the hot spot to the base of the transformer. That is,

$$T_h = T_c + T_1 + T_2 + T_3 \text{ etc.} \tag{9}$$

where

T_h = hot-spot temperature
T_c = base temperature

T_1, T_2, etc., are temperature rises in segments of the heat path. For any segment of the heat path

$$T = Kwd \tag{10}$$

where

w = density of heat flux in watts/square inches
d = length of segment in inches
K = a constant of thermal resistivity which has a value of 0.1 for copper, 0.2 to 0.4 for aluminum alloys, and about 200 for the best grades of solid insulation

Tables I, II, and III give the performance of the transformer shown in Fig. 5. This, although not an optimum design, demonstrates what can be accomplished.

A 3-phase core, solid-state rectifiers, and a bridge rectifier circuit also contributed to the size reduction of the magnetic components of this supply. The 3-phase core contributed about 15%, the bridge rectifier reduces the plate transformer size 25% and also reduces the filter choke size. The solid-state rectifiers make filament transformers unnecessary. Notice that over all, more than three-to-one weight reduction was accomplished.

Table II. Three-Phase Transformer Weight Comparison

Old Design		New Design	
Item	Weight, Pounds	Item	Weight, Pounds
Three single-phase transformers	7.50	3-phase trans- transformer	2.60
Choke	1.80	Choke	0.5
Filament transformer	0.33		
Total	9.63		3.1

Table III. Hot-Spot Temperature Rise of New 3-Phase Transformer

Condition	Degrees, C
With cold plate and natural convection	27
With cold plate and insulating cover to eliminate convection and radiation from transformer surface	33
With natural convection and radiation only	108
Calculated temperature rise assuming conventional design and natural convection and radiation only	127

Conclusions

It has been demonstrated that transformer size and weight can be greatly reduced. The best results will be obtained when all the factors affecting size and weight of the design are considered. Experience has shown that usually the transformer size can be reduced to less than one half the size now being used if a higher heat density can be allowed in the smaller transformer. This paper has shown how higher heat density can be used when better cooling or high-temperature insulation are employed. The techniques described here will make possible lighter air-borne electronic equipment without sacrificing performance or reliability.

Appendix

Loss Equations

If transformer size reduction is accomplished by increasing frequency while holding all other variables constant, then from equation 8

$$V_f = V_1/V = (f/f_1)^{3/4} = (f_r)^{-3/4} \tag{11}$$

where

V_f = the ratio of volume reduction
V = volume before reduction
V_1 = volume after reduction
f = frequency before size reduction
f_1 = frequency after size reduction
$f_r = f_1/f$ = ratio of frequency increase

Rearranging equation 11 becomes

$$f_r = (V_f)^{-4/3} \tag{12}$$

Similarly, when size reduction is accomplished by varying the other parameters one at a time.

$$(F_c)_r = (V_{Fc})^{-4/3} \tag{13}$$

$$B_r = (V_B)^{-4/3} \tag{14}$$

$$\Delta_r = (V_\Delta)^{-4/3} \tag{15}$$

When B and Δ are changed together and at the same rate $\Delta = kB$ where k is a constant. Substituting in equation 8

$$V_{\Delta B} = (B_s^2)^{-3/4} \tag{16}$$

Rearranging

$$B_s = (V_{\Delta B})^{-2/3} \tag{17}$$

where $B_s = B/B_1$ = ratio of flux increase = $\Delta/\Delta_1 = \Delta_s$ = ratio of current density increase, and $V_{\Delta B}$ = ratio of volume decrease when Δ and B are increased together at the same rate.

If an existing transformer design is considered in which core loss = coil loss, then as size is being reduced:

$$\text{Loss ratio} = \frac{\text{core loss ratio} + \text{coil loss ratio}}{2} \tag{18}$$

As the frequency is increased the loss per cubic inch of iron increases as the 1.43 power of frequency. Loss per cubic inch of conductor remains constant since the current density remains constant. Then

$$L_f = \frac{V_f(f_r)^{1.43} + V_f}{2} \tag{19}$$

Substituting from equation 12

$$L_f = \frac{(V_f)^{-0.91} + V_f}{2} \tag{20}$$

where

L_f = the ratio of loss increase as the transformer size is being reduced by increasing frequency while holding all other parameters constant

When the size reduction is accomplished by increasing F_c the loss per cubic inch of conductor and the loss per cubic inch of core stay constant, but the ratio of conductor volume is $(F_c)_r(V_{Fc})$.

$$L_{Fc} = \frac{(F_c)_r(V_{Fc}) + V_{Fc}}{2} \tag{21}$$

Substituting from equation 13 and simplifying

$$L_{Fc} = \frac{(V_{Fc})^{-1/3} + V_{Fc}}{2} \tag{22}$$

where L_{Fc} is loss ratio as the transformer size is being reduced by increasing coil space factor.

When the size reduction is accomplished by increasing current density Δ, the loss per cubic inch of conductor increases as the square of current density while the core loss per cubic inch stays constant.

$$L_\Delta = \frac{V_\Delta + V_\Delta(\Delta_r)^2}{2} \tag{23}$$

Substituting from equation 15 and simplifying

$$L_\Delta = \frac{V_\Delta + (V_\Delta)^{-5/3}}{2} \tag{24}$$

where

L_Δ = loss ratio as size is decreased by increasing Δ

The situation is the same when B is increased except that coil loss and core loss are interchanged:

$$L_B = \frac{V_B^{-5/3} + V_B}{2} \tag{25}$$

When size reduction is achieved by increasing B and Δ at the same rate, the loss per cubic inch of core increases as B^2 and the loss per cubic inch of coil varies as Δ^2.

$$L_{\Delta B}=\frac{(V_{\Delta B})(B_s)^2+(V_{\Delta B})(\Delta_s)^2}{2} \qquad (26)$$

Substituting for equation 17 and simplifying

$$L_{\Delta B}=(V_{\Delta B})^{-1/3} \qquad (27)$$

where

$L_{\Delta B}$ = the loss ratio when transformer size is reduced by increasing both flux density and current density at the same rate

Heat Density Equations

The heat density ratio is the loss ratio divided by the volume. When size reduction is accomplished by increasing frequency.

$$H_f=\frac{L_f}{V_f} \qquad (28)$$

where

H_f = ratio of heat dissipated per cubic inch after size reduction to heat dissipated, per cubic inch, before size reduction

The same is true for other parameters. For instance:

$$H_B=\frac{L_B}{V_B} \qquad (29)$$

Temperature Rise Equations

Assume that a transformer is to be reduced in size while total heat dissipated is to remain constant. Assume that in this transformer one half the temperature rise above ambient is at the surface of the transformer and that the other half of the temperature rise is due to temperature gradient within the transformer. Assume also that as the transformer size is reduced the relative heat distribution, thermal conductivity, and thermal emissivity remain constant. Then:

$$T_r=\frac{T_s+T_g}{2} \qquad (30)$$

where

T_r = ratio of the total temperature rise in the transformer of reduced size to the total in the transformer before size reduction

T_s = ratio of temperature rise at the surface after size reduction to the rise at the surface before size reduction

T_g = ratio of temperature rise inside the transformer after size reduction to the temperature rise before size reduction

$$T_s=\frac{1}{\text{ratio of surface area}}$$

$$=\frac{1}{(V_r)^{2/3}}=V_r^{-2/3} \qquad (31)$$

where

V_r = the ratio of volume after size reduction to the volume before size reduction

T_g = ratio of heat flow density × ratio of heat flow path length (32)

$$T_g=(V_r)^{-2/3}\times(V_r)^{1/3}=V_r^{-1/3} \qquad (33)$$

Substituting equations 31 and 33 in equation 30

$$T_r=\frac{(V_r)^{-2/3}+(V_r)^{-1/3}}{2} \qquad (34)$$

When size reduction is accomplished by increasing B and Δ at the same rate, the relative heat distribution within the transformer remains constant. Then

$$T_{\Delta B}=L_{\Delta B}T_r \qquad (35)$$

where

$T_{\Delta B}$ = temperature rise ratio in a transformer reduced by increasing B and Δ at the same rate

Combining equations 27 and 35

$$T_{\Delta B}=\frac{(V_{\Delta B})^{-1}+(V_{\Delta B})^{-2/3}}{2} \qquad (36)$$

References

1. KEEPING AIRCRAFT WEIGHT AND COST DOWN, **E. E. Heinemann.** *Journal*, Society of Automotive Engineers, Lancaster, Pa., vol. 63, no. 4, Apr. 1955, pp. 25–28.

2. RESEARCH AND DEVELOPMENT OF NEW DESIGN METHODS FOR POWER TRANSFORMERS, **G. A. Forster, L. J. Stratton, H. L. Garbarino.** *Final report,* Electronics Parts and Materials Branch, Signal Corps Engineering Laboratories, Armour Research Foundation, Illinois Institute of Technology, Chicago, Ill., Mar. 1956, pp. 1–17.

A Digital Flying Extensometer for Temper Rolling Mills

N. S. WELLS
MEMBER AIEE

STEEL STRIP is extremely hard after it has been cold rolled and it must then be softened by annealing. The next production process in the manufacture of tinplate is temper rolling. Temper mills are used to produce the desired surface finish in the strip with acceptable flatness plus the desired mechanical properties. The latter refers mainly to the elimination of a surface defect known as "stretcher-strain" markings which can occur in pressed-steel products. The tendency to this type of defect is directly related to the amount of yield-point elongation. If the steel can be subjected to a small amount of cold working this defect can be eliminated. The effect of various amounts of cold reduction upon the yield-point stress and the stress-strain relationship is shown in Fig. 1. It will be noted that with an extension of as little as 1.0% the yield point can be entirely suppressed.

This extension is achieved in a temper mill by a combination of strip tension and roll pressure in the two mills. Fig. 2 shows the pass line arrangement for such a mill. Certain of these mills can roll at strip speeds of 6,000 fpm (feet per minute). In all sections of the mill the strip is under substantial tension.

It is standard mill production practice to check the amount of work hardening produced in the mill by means of Rockwell hardness tests. Steel samples are taken from each coil, on the strip at the entry and exit sides of the mill. This is a spot check and can only be carried out by stopping the mill. One other method of checking strip elongation is by scribing the strip on the pay-off reel and rechecking the same scribe marks on the wind-up reel. Here a scribe gauge length of $12^1/_2$ inches is used resulting in a 1% extension figure for each 1/8-inch addition measured on the outgoing side of the mill. This scribing technique can only be carried out at a mill speed of around 200 fpm on a normal production run basis. For special tests however, the coil is run slowly through a processing line and scribed throughout its length. After temper rolling the coil is unwound and the elongation checked. This is a very costly and time consuming technique. There is therefore a definite need for an extens-

---

Paper 57-703, recommended by the AIEE Industrial Control Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted April 18, 1957; made available for printing June 14, 1957.

N. S. WELLS is with The Steel Company of Canada, Ltd., Hamilton, Ont., Canada.

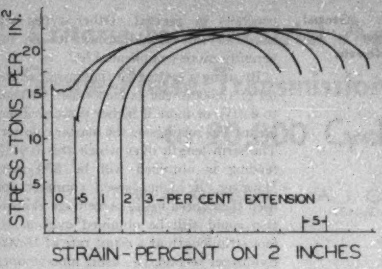
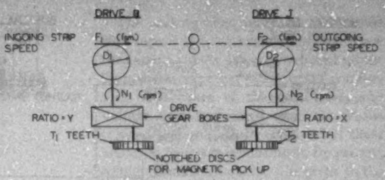


Fig. 1 (left). The influence of cold rolling on the shape of the stress-strain curve of annealed low carbon steel

Fig. 3 (right). Details of drives B and J



ometer of suitable design which will permit a continuous observation and recording of strip elongation under all mill running conditions.

Design Basis of Digital Extensometer

The strip speed at entry and exit ends of the mill is assumed to be identical to the peripheral speed of the bottom entry tension and bottom exit tension rolls respectively. This is a reasonable assumption since in both of these drives the strip is in close contact with a considerable proportion of the roll circumference and is subjected to both front and back tension.

The percentage strip elongation E can be expressed as (Fig. 3):

$$E=\frac{F_2-F_1}{F_1}100\%$$

$$\text{or } \frac{N_2D_2-N_1D_1}{N_1D_1}100\%$$

If P_1 = pulses per minute detected by entry magnetic pick-up unit

Then

$$P_1=N_1T_1Y$$

Similarly

$$P_2=N_2T_2X$$

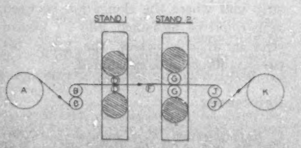


Fig. 2. Two-stand temper mill pass line arrangement

A—Payoff rolls
B—Tension rolls
D—Twelve-inch-diameter work rolls
F—Tensiometer roll
G—Twenty-inch-diameter work rolls
J—Exit tension rolls
K—Wind-up roll

$$E=\frac{KP_2-P_1}{P_1}100\%$$

$$\text{where } K=\text{constant}=\frac{D_2T_1Y}{D_1T_2X}$$

Let C_2 be the total counts received at J when C_1 counts are received at B, i.e., considering a per unit time basis.

If $C_1=1{,}000K$

Then

$$E=\frac{C_2-1{,}000}{1{,}000}100\%$$

This equation can be solved directly using a pair of multidecade counters.

One of these counters receives pulses from drive B and is designated the "master" unit; the other is pulsed from drive J and is the "slave" unit. The master unit is preset for a total pulse count of 1,000K. Immediately coincidence is reached between preset and actual counts a blocking pulse switches off the pulse train to the slave counter. Then for a gated count of 1,000 on the slave unit there will be zero elongation in the strip, with the master counter preset for the correct value of 1,000K.

1. For a count of 1,010, the elongation will be 1%.
2. For a count of 1,025, the elongation will be 2.5%.

The precise value of strip elongation from zero to 9.9% can therefore be directly obtained by reading the last two decades of the slave counter.

The accuracy of the results obtained is determined entirely by the accuracy of the calculated constant K. If a value $K^1=aK$ is used where K is the correct constant value and a is a factor which is not necessarily unity, the preset master count will be 1,000K^1, and

$$E=\frac{KC_2-1{,}000K^1}{1{,}000K^1}100\%$$

$$=\frac{C_2-1{,}000a}{1{,}000a}100\%$$

Fig. 4 shows the resulting effects on calibration for various values of a. This suggests an alternative method of calibrating the extensometer using a prescribed coil and plotting actual and indicated elongations. The line $a=1.005$ is an example of some results obtained from an actual test. In this particular case the original constant was calculated to be 968 using gear ratios instead of gear tooth numbers. The value of K indicated by the curve is 968/1.005 = 963, which agreed with the recalculated constant using gear tooth numbers.

Equipment Details

The general arrangement of the units comprising this extensometer is shown in Fig. 5. Noncontact magnetic pick-ups provide a simple method of counting pulses corresponding to roll revolutions, although there is the slight complication involved in mounting suitable toothed wheels. It is possible to use the actual mill rolls as the pulse producing units. In this case the outer edge of the roll has equally spaced slots milled around it. These slots need only extend a very small distance axially. This does not interfere in any way with the strip in the mill since the full face width of any roll is never fully utilized during rolling. The magnetic pick-up unit is then mounted below the roll to avoid damage to it during steel wrecks in the mill. In the 2-stand temper

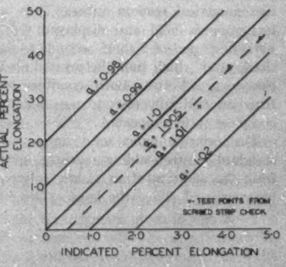


Fig. 4. Calibration curves for various values of a

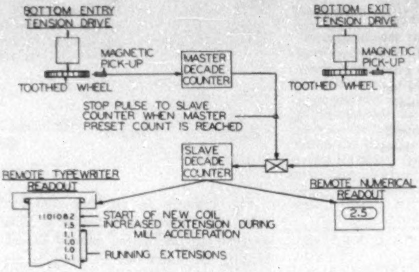

BOTTOM ENTRY TENSION DRIVE
MAGNETIC PICK-UP
TOOTHED WHEEL
MASTER DECADE COUNTER
BOTTOM EXIT TENSION DRIVE
MAGNETIC PICK-UP
TOOTHED WHEEL
STOP PULSE TO SLAVE COUNTER WHEN MASTER PRESET COUNT IS REACHED
SLAVE DECADE COUNTER
REMOTE TYPEWRITER READOUT
1101082
1.5
1.1
1.0
1.0
1.1
START OF NEW COIL
INCREASED EXTENSION DURING MILL ACCELERATION
RUNNING EXTENSIONS
REMOTE NUMERICAL READOUT
2.5

Fig. 5. General arrangement of extensometer

mill to which this digital extensometer has been applied, the tensiometer roll *F* was modified in this fashion and operated satisfactorily. Using this technique it is possible to measure strip elongation on single-stand temper mills.

The particular type of magnetic units used will not follow a pulse count rate less than that corresponding to approximately 200 fpm strip speed. This is not particularly significant as this corresponds to thread speed on the mill, i.e., not a true rolling condition. The tooth numbers on the counting wheels are arranged to give a maximum count rate of 1,500 per second with the mill running at a speed of 3,000 fpm. This corresponds to a measurement of elongation on a 30-foot section of strip approximately.

Two identical 5-decade decimal counters utilizing cold-cathode glow-transfer counting tubes are used with maximum count rates of 4,000 per second. The master counter incorporates an electronic switching unit which blocks the pulse signals into the slave counter when the master preset count is reached. This preset count, i.e., $C_1 = 1{,}000K$ is set up on standard 10-position switches available on each decade. The slave counter is so arranged that remote indication is provided for the last two digits utilizing two numerical readout tubes. This information is held and displayed for a variable period of 2 to 15 seconds as desired, after which time the printed readout is typed out. Both counters are then automatically reset to zero and the complete cycle is repeated.

The print-out unit used employs a standard electric adding machine apart from the addition of an 11-key solenoid box mounted above the keys. The 60-cycle supply to the printer is removed by means of a limit switch on the main mill speed setting rheostat for strip speeds of less than 300 fpm. The solenoid keys can still be operated by the counter, but since a mechanical column transfer is used a seven figure print-out is obtained for the first reading after the supply has been restored. Identification of each new coil is readily obtained in this way. Fig. 5 indicates the type of record produced.

In this particular application the counter rack is mounted in the motor room adjacent to the mill. The numerical decimal read-out unit displaying strip extension is located at the rollers control position and the print-out unit is mounted in a suitable enclosure adjacent to the mill recorder. A test switch on the roller's visual read-out unit enables a check to be made of the two counters plus the slave pulse blocking circuit. This is achieved by switching the slave counter to a second magnetic pick-up unit on drive *B*. If everything is functioning correctly the two counters will always show the same count, which will be the master preset count. The roller's unit similarly will always show the same two digits.

Conclusion

This digital extensometer is basically simple, the quantity of equipment involved is small and relatively cheap. There are few if any installation problems and calibration can be readily checked. This general type of equipment involving the numerical presentation of information will become general in steel plant operations in particular, and industrial concerns in general. Other types are available but would appear to be inherently more complicated.[1,2]

By using a setting on the master counter of 10,000*K* the extension can be read to 0.01% or more if higher powers of ten are used as settings on the master counter. The strip length over which this 0.01% reading is obtained will be 300 feet. However, if higher speed counters are used then three figure percentage elongation could still be obtained over a 30-foot strip length at a count rate of 10,000 counts per second, i.e., using more counts per foot of strip. If more counts are made per foot of strip but the 1,000*K* setting is used on the master counter, elongation will then be measured over 3 feet of strip and a reading will be obtained in 1/10 second approximately.

It then becomes feasible to consider a closed-loop-type of control system utilizing this digital extensometer information as the feedback in a sampling type of servo system and controlling the strip elongation in the mill to any desired set value. Results obtained on this particular temper mill indicate that running conditions obtained are very stable and the strip elongation remains essentially constant. Extension does increase during acceleration and deceleration—indicating some inaccuracy in the setting of mill inertia compensation circuits. Automatic extension control would not appear to be justified in this particular case.

However the use of this extensometer for closed loop control in other manufacturing processes can be considered. One application would be on paper machine drives where extremely accurate control of the speed difference between two adjacent stands is necessary. Here the digital extensometer could perhaps be used as a vernier-type sampling servo control in conjunction with a main section voltage control. Another measurement application would be on a multistand hot strip mill where the elongation between mills could be measured, and for known steel input thicknesses, the gauge between mills could be calculated.

References

1. Extensometer. *U. S. patent nos. 2,447,208; 2,447,209* and *2,474,116*, United States Steel Company, Pittsburgh, Pa.

2. Baldwin Magnetic Extensometer. *Report no. MW/A/34/54*, British Iron and Steel Research Association, London, England, 1954.

Criteria for Industrial Application of Single-Phase Transmission Lines, 400 to 20,000 Cycles

W. A. MUNSON
ASSOCIATE MEMBER AIEE

A. J. GERMAIN
NONMEMBER AIEE

SINCE the early days of alternating current, electrical engineers have designed many forms of generating devices to enable them to study the effects of different frequencies on an electric circuit or machine. This study has been a continuing one due to the foreseeable advantages of high-frequency power in certain areas.

During the past decade, the use of audio frequencies has been increasing. This increase has not been limited to merely a higher kilowatt consumption, but has also taken the form of an increased number of industrial and commercial applications. Also, the increased use of audio frequencies has resulted in different methods of generation of these frequencies.

The use of audio-frequency (a-f) power is a recognized power source, today, for such various functions as internal grinding, portable tools for high-production assembly lines, woodworking machines, industrial and commercial high-efficiency fluorescent lighting, induction heating, electroluminescence panels, and ultrasonic applications.

Table I lists the known applications and the range of frequencies presently used for the particular application.

Methods of Generation

The methods of generating audio frequencies have kept pace somewhat with the use of audio frequencies, thus resulting in several generating devices that are now fully developed and a few newer types that are still in the developmental stage. This parallel development of generation with application has provided the impetus for some of the newer higher frequency applications.

Some of the fully developed methods are the motor-generator combinations, the electronic tube oscillator, and the spark-gap-type generator. These devices are used over a wide range of frequencies, and when coupled with frequency multipliers, are capable of producing the ultrasound frequency, 20,000 cps (cycles per second).

The motor-generator combinations, when fully developed, use a conventional motor, usually 60-cycle a-c, driving an inductor or salient-pole generator. The newer trend today is to the use of a permanent-magnet-type generator rather than those previously mentioned. However, this newer form of generator has not yet been fully developed and as such its application is still rather limited.

The new transistor oscillator is now in a full-scale development program as it is anticipated that this new form of generation will overcome many of the shortcomings of audio-frequency present generation schemes.

The normal frequencies which have been standardized on and are commercially available are 400–500 cps, 1,000 cps, 3,000 cps, 10,000 cps, and 20,000 cps. On some of the higher frequencies, 10,000 and 20,000 cps, the required frequency may be obtained through the use of a generator producing a lower frequency coupled with a frequency doubler or tripler.

Audio-Frequency Transmission

As is the case with direct current and the conventional power frequencies, 60 cps and below, audio frequencies also require the use of metallic conductors to transmit their power to remote locations. However, unlike direct current or the conventional power frequencies, great consideration must be given to the transmission system to obtain efficient transmission of the a-f power.

The factors that must be considered with regard to the transmission circuit are the same as those for conventional power frequencies: current rating of the conductors, inductance of the conductor, and voltage and/or power loss in the transmission system. However, as all of these elements bear a relationship to frequency, minute consideration must be given to the system to obtain efficient transmission at the relatively high frequencies in question, the audio frequencies.

Fig. 1 shows the maximum allowable reactance, in ohms, which can be tolerated for various loads without leaving a rise in voltage at the point of load utilization greater than 5% after correcting the generator load to unity power factor through the use of capacitors at the load utilization point. The voltage rise of 5% is based on a nonfunctioning load and a fully loaded generator. As most a-f generators are designed to supply only a real power load. kilowatts only, the necessity for transmission system and load power factor correction is obvious. Where load fluctuations will be present, the over-all complete system must be analyzed for all possible load conditions to prevent unduly high overvoltages at the load when the generator is not fully loaded.

Thus, any transmission system selected should not exceed the maximum allowable line reactance values given in Fig. 1 if satisfactory and efficient a-f power transmission is to be obtained.

It is desirable to use standard cable if at all possible, both from the standpoint of cost and availability. Table II gives the properties for various cable combinations as shown in Fig. 2.[1] This table gives cable ratings for the commercially available standard audio frequencies. Below 400 cycles, standard 60-cycle cable data are used as a basis for choosing a transmission system.

Fig. 2 shows the arrangement of the cables as used in Table II, and also the current distribution that can be anticipated in the particular cable arrangement selected. The table gives not only the anticipated ohms but also the allowable current capacity of the particular cable arrangement as well as its expected wattage loss. All values given are on a per-foot basis and include both conductors for a single-phase 2-conductor system.

Tables III and IV[2] give characteristics of electric cables of the type usually used in the arrangement given in Fig. 2. The use of these two tables in conjunction with Fig. 2 will enable the determination of the over-all size of the transmission system.

The selection of a transmission system conductor size becomes simple by merely

---

Paper 57-1021, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 7–11, 1957. Manuscript submitted June 10, 1957; made available for printing August 13, 1957.

W. A. MUNSON and A. J. GERMAIN are both with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

The authors wish to acknowledge the assistance of members of the engineering department of the Westinghouse Industrial Electronics Division in the preparation of many of the tables included in this paper, and to express their appreciation to the Ladish Company for permission to discuss the installed high-frequency system.

Table I. Applications and Associated Frequencies

Frequency, Cycles	Application
180	Resistance welding, continuous tube mills
180–360	High-cycle tools
180–420	High-speed grinders
180–720	High-speed woodworking machines
360–500	Ozonators
360–10,000	High-efficiency fluorescent lighting*
400–500	SONAR
400, 600, 800	Aviation industry
500	Alternating current to high-voltage direct current, 20 kv to 1,500 kv (for cable testing)
600, 800	Navy
960, 3,000, 4,320, 6,600, 10,000	Induction heating
3,000	Electroluminescence panels
20,000	Ultrasound applications

* 840 cycles is the most popular frequency at the present time for high-efficiency fluorescent lighting.

determining what current must be handled, and then making certain that the conductor reactance is not greater than the allowable maximum value as determined from Fig. 1. The third feature which must be checked is the power loss in the transmission system. For practicability, this is usually limited to 5% of the load rating. Here the 5% value is a rule-of-thumb feature and must be balanced against the increased cost of the transmission system to reduce the power loss to 5% assuming the other features, current and maximum allowable transmission reactance ohms, have been met. However, experience has indicated that a maximum loss of more than 5% is normally very uneconomical.

The presently preferred system of the various arrangements shown in Fig. 2 is the one where four conductors are used with similar poloraties opposite. This arrangement has had most success due to its relatively low value of reactance. Several large installations consisting of several generators paralleled at a central location and transmitting the power then for several hundred feet use this 4-conductor method.

Fig. 3 shows a comparison between the three systems outlined in Fig. 2. This comparison is between copper area, reactance, and resistance for various cable combinations, all of which are capable of handling the same amount of a-f current and/or power.

Future Trends

As the use of a-f power increases, the requirement of using larger, centrally located generators becomes an economical requirement. The problem of transmitting these larger blocks of power to a point of utilization has been the subject of tremendous activity.

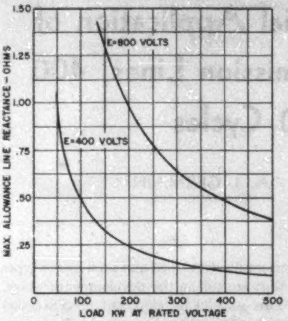


Fig. 1. Maximum allowable line reactance based on ability to tune generator load to unity power factor with 5% rise in voltage at load utilization

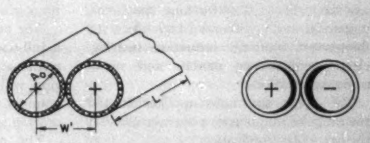
The coaxial system, used for many years for frequencies well above the audio range, has been vigorously investigated. The adaptability of the type of system has been enhanced through the recent development of a portable water-cooled coaxial conductor. However, such a system lends itself to relatively short runs which is not a desired feature.

Another type of transmission system, known as the concentric conductor system, has been in service for several years at the Ladish Company Forge Plant in Cudahy, Wis. This system which has been patented by Ladish has enjoyed a degree of success equal to its original expectations.

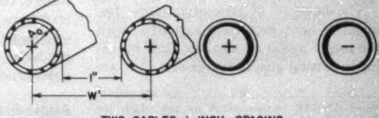
The concentric conductor system is similar to a coaxial system in that the two conductors are co-axially formed. However, whereas the inner conductor is located at the center of the cable for the coaxial system, the inner conductor of the concentric conductor system is located as tight against the outer conductor as insulation will permit. Fig. 4 shows the major difference between a coaxial system and a concentric conductor system.

The system presently in use at the afore-mentioned plant transmits 3,600 cps power at 4,800 volts which is then stepped down to either 200 or 400 volts at the various utilization points. At the present time, two 1,750-kw 3,600-cps generators are in service with a third 1,750-kw generator presently being designed to parallel with the two existing units. This third unit, when installed, will result in a total installed capacity of 5,250 kw of 3,600 cps power.

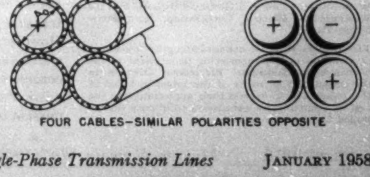
The 3,600-cycle power is distributed throughout the plant through the use of 4 and 2-inch concentric conductor transmission lines with 2-inch concentric conductors used for power take-offs and



TWO CABLES—SIDE-BY-SIDE



TWO CABLES—I INCH SPACING



FOUR CABLES—SIMILAR POLARITIES OPPOSITE

Fig. 2. Typical conductor arrangements and current distribution in the conductors

Table II. Resistance and Impedance Values for Cables in Free Air: Extra Flexible Stranded Cable at 30 C Ambient Temperature,* 85 C Maximum Copper Temperature

Cable Size, B&S or MCM†	Two Cables, Side by Side‡				Two Cables, 1-Inch Spacing‡				Four Cables, Similar Polarities Opposite‡			
	Watts Per Foot Run	Current Rating, Amperes	Resistance Milliohms Per Foot	Impedance, Milliohms Per Foot	Watts Per Foot Run	Current Rating, Amperes	Resistance, Milliohms Per Foot	Impedance, Milliohms Per Foot	Watts Per Foot Run	Current Rating, Amperes	Resistance, Milliohms Per Foot	Impedance, Milliohms Per Foot
400–500 Cycles												
14	8.0	34	6.7	6.7	8.0	34	6.8	7.0	11.9	62	3.4	3.4
10	8.7	62	2.45	2.6	8.7	63	2.43	2.25	12.5	105	1.2	1.28
6	11.6	98	1.2	1.18	12.0	110	1.03	1.35	14.8	180	0.47	0.47
4	11.8	125	0.7	0.72	12.2	140	0.61	0.95	14.9	230	0.29	0.30
2	12.0	173	0.45	0.62	12.7	190	0.33	0.82	15.0	295	0.19	0.255
1/0	13.6	240	0.35	0.53	14.0	270	0.187	0.75	17.8	410	0.11	0.22
4/0	15.2	290	0.3	0.51	16.3	345	0.13	0.67	20.0	520	0.075	0.19
250	17.1	330	0.22	0.48	18.8	395	0.12	0.63	22.1	590	0.065	0.19
350	18.9	370	0.18	0.45	20.4	450	0.10	0.59	24.6	670	0.054	0.17
500	20.8	440	0.15	0.40	23.0	600	0.0775	0.54	26.8	780	0.045	0.16
1,000 Cycles												
14	8.0	34	6.90	7.0	8.0	34	6.82	7.57	11.6	58	3.45	3.5
10	8.7	59	2.50	2.82	8.7	60	2.40	3.50	12.5	100	1.25	1.41
6	11.6	98	1.21	1.76	12.2	107	1.06	2.52	14.8	176	0.484	0.704
4	11.3	123	0.75	1.37	12.4	140	0.630	2.18	14.9	224	0.30	0.548
2	11.5	151	0.507	1.17	12.6	175	0.410	1.88	15.0	272	0.203	0.468
1/0	13.6	197	0.350	1.025	13.1	234	0.270	1.71	17.8	355	0.140	0.410
4/0	15.2	240	0.265	0.915	16.7	296	0.190	1.52	20.0	433	0.106	0.366
250	16.9	270	0.234	0.890	19.0	328	0.176	1.47	22.1	487	0.093	0.356
350	18.9	305	0.203	0.816	20.3	372	0.147	1.35	24.6	550	0.081	0.326
500	20.6	358	0.160	0.752	24.6	463	0.115	1.25	26.6	645	0.064	0.301
3,000 Cycles												
14	7.9	33	7.25	8.06	7.9	34	6.86	11.3	11.3	56	3.62	4.03
10	8.5	56	2.70	4.82	8.6	59	2.46	8.06	12.2	95	1.35	2.41
6	11.6	96	1.26	3.94	12.0	105	1.10	6.85	14.8	172	0.504	1.58
4	11.9	117	0.87	3.54	12.2	129	0.735	6.30	15.6	212	0.348	1.41
2	12.0	130	0.71	3.10	12.7	148	0.580	5.40	15.7	235	0.284	1.24
1/0	13.8	158	0.552	2.72	14.6	185	0.426	4.94	18.0	285	0.221	1.09
4/0	15.2	187	0.434	2.40	16.1	224	0.320	4.40	19.8	338	0.173	0.96
250	17.2	211	0.384	2.38	18.9	255	0.291	4.30	22.5	382	0.154	0.95
350	18.2	236	0.325	2.17	20.4	296	0.233	3.90	23.5	425	0.130	0.87
500	20.8	278	0.270	2.04	22.5	340	0.195	3.60	27.0	500	0.108	0.82
10,000 Cycles												
14	8.1	33	7.44	16.50	8.1	34	6.99	29.80	13.4	60	3.72	8.25
10	8.75	53	3.10	13.30	9.0	56	2.85	25.30	14.0	95	1.55	6.65
6	11.0	76	1.93	11.95	11.6	82	1.71	22.13	14.5	137	0.772	4.78
4	11.8	91	1.42	10.50	12.0	100	1.205	20.00	15.2	164	0.568	4.20
2	12.3	100	1.23	10.05	12.8	113	1.000	17.40	15.9	180	0.490	4.03
1/0	13.5	117	0.99	8.29	15.0	139	0.775	15.91	17.3	210	0.395	3.31
4/0	15.3	145	0.725	7.64	16.6	176	0.533	14.10	20.0	262	0.290	3.06
250	17.2	162	0.655	7.57	18.6	194	0.495	13.70	22.2	292	0.262	3.03
350	19.5	184	0.575	6.71	20.4	222	0.415	12.67	25.3	332	0.230	2.69
500	21.2	208	0.480	6.41	22.7	258	0.340	11.83	27.1	375	0.192	2.56
20,000 Cycles												
14	8	32	7.8	30.0	8.1	34	7.0	57.5	13.0	57.6	3.9	12.8
10	8.7	49.5	3.58	25.45	9.1	51.9	3.4	49.5	14.0	84	2.0	11.3
6	11.6	66.2	2.65	23.2	11.0	66.5	2.5	43.7	14.8	117	1.08	8.3
4	11.9	74.2	2.17	20.3	11.7	76.6	2.0	39.5	15.6	135	0.865	7.4
2	12.4	79.8	1.96	19.8	13.0	96.5	1.6	34.6	16.0	144	0.765	7.05
1/0	13.6	99.4	1.38	16.2	15.0	109	1.26	31.5	17.8	168	0.63	5.7
4/0	15.4	120.0	1.065	15.1	16.6	142	0.82	28.5	20.0	208	0.46	5.3
250	17.2	132.0	0.980	14.9	19.0	161	0.78	27.0	22.5	238	0.4	5.25
350	20.1	151.0	0.885	13.47	24.0	190	0.67	25.5	25.5	266	0.36	4.6
500	21.7	176.0	0.702	12.7	22.7	205	0.54	23.3	27.1	296	0.31	4.35

* Multiply ampere rating by following value b for rating at other than 30 C (degrees centigrade) ambient: at 20 C, by 1.08; at 40 C, by 0.90; at 50 C, by 0.79.

† Brown and Sharpe gage or thousand circular mils.

‡ For cables in nonmagnetic conduit, multiply current values by 0.73.

Table III. Characteristics of Electric Cable: Extra Flexible, Varnished Cambric Insulation, Glyptal Treated Braid Cover, 85 C Maximum Copper Temperature

Size, B&S or Closest MCM*	Conductor			Diameter in Inches		Weight Per 1,000 Feet, Pounds
	Number of Strands	Size of Strands	Insulated Thickness in 64ths	Conductor	Over-All	
14	19	27	2	0.071	0.185	24
10	27	24	3	0.121	0.265	65
6	61	24	4	0.181	0.370	156
4	125	24	4	0.261	0.455	268
2	175	24	5	0.360	0.585	361
0	275	24	5	0.460	0.685	532
0000	550	24	5	0.613	0.855	976
250	650	24	6	0.673	0.945	1190
350	925	24	6	0.799	1.070	1642
500	1325	24	7	0.970	1.275	2345

* Brown and Sharpe gage or thousand circular mils.

Sable IV. Characteristics of Electric Cable: Standard Strand, Varnished Cambric Insulated, Single Braided, 85 C Maximum Copper Temperature

Size, B&S or Closest MCM*	Conductor			Diameter in Inches		Weight Per 1,000 Feet, Pounds
	Number of Strands	Diameter of Each Strand, Inches	Insulated Thickness in 64ths	Conductor	Over-all	
6	7	0.0612	4	0.184	0.373	130
4	7	0.0772	4	0.232	0.422	180
2	7	0.0974	4	0.292	0.482	265
0	19	0.0745	5	0.373	0.594	415
0000	19	0.1055	5	0.528	0.749	765
250	37	0.0822	6	0.575	0.828	908
350	37	0.0973	6	0.681	0.955	1260
500	37	0 1162	6	0.814	1.172	1866

* Brown and Sharpe gage or thousand circular mils.

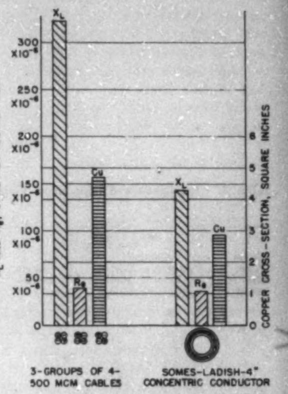


Fig. 5. Comparison of physical characteristics for a 4-cable conductor arrangement and the Somes-Ladish 4-inch concentric conductor

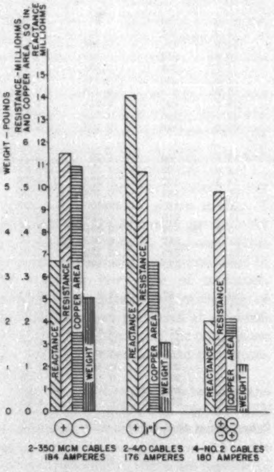


Fig. 3. Comparison of various conductor arrangements, each capable of handling 175 amperes at 10,000 cps

secondary distribution when required. The total length of the high-voltage (4,800-volt) distribution system is now equal to approximately 1,000 feet of 4-inch concentric conductors and 2,000 feet of the 2-inch concentric conductor. In addition, there is approximately 1,000 feet of 4-inch and 2-inch concentric conductors used for the low-voltage distribution.

The concentric conductor system has many advantages of which the following are the major ones:

1. Rigid all-metallic conductor system.
2. Grounded outer shell.
3. Very desirable electrical characteristics; i.e., low reactance, low resistance, reduced copper cross section.
4. Complete selection of power take-off fittings and expansion points.
5. Increased current capacity, if necessary, by gas or liquid cooling.
6. Relative simplicity coupled with low installation costs.

Fig. 5 shows the comparison between a concentric-conductor and a 4-conductor line, similar polarities opposite, for resistance, reactance, and copper cross section

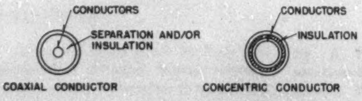


Fig. 4. Comparison between a conventional coaxial conductor and a concentric conductor

Conclusions

The information given in Tables II–IV used in conjunction with Fig. 1 will enable the engineer, who is faced with the necessity of designing an a-f transmission system, to determine what size and/or type of conductor arrangement to use. Also, the information given will enable him to allow for the required amount of space for installation of an a-f transmission system. At all times, however, it must be remembered that the system, when once selected, is not to be placed in magnetic conduit as the frequencies in question, 400 cps up to and including 20,000 cps, would induce excessive eddy current losses in conventional magnetic material conduit. Thus, for all a-f transmission systems, the conductors, when necessary to be enclosed in conduit, must be enclosed by nonmagnetic material.

References

1. Characteristics of Single-Conductor Electric Cable at High Frequency, J. T. Sabol. *AIEE Transactions*, vol. 74, pt. III, 1955, pp. 1280–85.

2. Industrial Electronics Reference Book (book), Westinghouse Electronics Engineers. John Wiley & Sons, Inc., New York, N. Y., 1948, pp. 386–87.

3. Induction Heating (book), N. R. Stansel. McGraw-Hill Book Company, Inc., New York, N. Y., 1949, chap. 8.

4. Inductance Calculations (book), F. W. Grover. D. Van Nostrand Company, Inc., Princeton, N. J., 1946, chap. 24.

5. STANDARD HANDBOOK FOR ELECTRICAL ENGINEERS (book), A. E. Knowlton. McGraw-Hill Book Company, Inc., New York, N. Y., eighth edition, 1949, sects. 12–476, 12–477, 15–112.

6. IMPEDANCE DATA FOR 400-CYCLE AIRCRAFT DISTRIBUTION SYSTEMS, D. W. Exner, G. H. Singer, Jr. *AIEE Transactions*, vol. 71, pt. II, 1952, pp. 410–19.

7. UTILIZATION OF 300 TO 2,000 CYCLE POWER IN INDUSTRY, V. C. Geckler. *Proceedings*, American Power Conference, Chicago, Ill., vol. 18, 1956, pp. 482–91.

The Selection and Use of Servo Performance Criteria

W. C. SCHULTZ
ASSOCIATE MEMBER AIEE

V. C. RIDEOUT
MEMBER AIEE

THE APPLICATION of performance criteria in the analysis, and more particularly in the synthesis, of servo systems is a matter of importance, but it is one of considerable difficulty as compared with the application of stability criteria. Although it is generally agreed that concise criteria of performance are desirable, there has been no general agreement as to what criteria should be used. This is not too surprising, for surely there should be many criteria to fit the various requirements of different servos.

Perhaps the simplest mathematical statement that can be made regarding servo performance is the error, $e(t)$, defined as the difference between the input and output,

$$e(t)=r(t)-c(t) \tag{1}$$

For comparison purposes, however, such a function of time is not a convenient expression, since, in general, it is hard to define just exactly what is meant by "comparison" of two functions. Concepts that might be borrowed from the mathematician here include "approximation in the mean," "uniform convergence," the value of the maximum absolute difference, and the like. The underlying difficulty encountered in the establishment of a general performance criterion is the choice of the best concept to use in any particular case. One very reasonable procedure is to choose, as a basis for a general criterion, a functional,[1] that is to say, a number, which is determined by the whole range of any one of a class of functions.* Although, to the authors' knowledge, no general theory has been set forth which proposes functionals as a means of stating performance criteria, there have been various proposals and suggestions of specific cases in the literature.

If $r(t)$ and $c(t)$ are stationary random processes, the ordinary mean-square measure of error,

$$E=\lim_{T->\infty}\frac{1}{2T}\int_{-T}^{T}e^2(t)\,dt \tag{2}$$

is often used. In this paper, only transient input functions will be considered. In such cases, as discussed by Hall,[2] a measure proportional to the mean-square error may be written as

$$E=\int_0^\infty e^2(t)\,dt \tag{3}$$

An extension of equation 2 to the case in which delay (or advance) of the output is considered has been given by Spooner and Rideout.[3] Their form, called the generalized error function (GEF), may also be adapted to transient inputs and will be discussed in more detail in a subsequent section of this paper.

A simpler measure of error, proposed by Oldenbourg and Sartorious,[4] is

$$E=\int_0^\infty e(t)\,dt \tag{4}$$

A number of other measures of error have been proposed[5–10] and used, including the integral of the absolute error,

$$E=\int_0^\infty |e(t)|\,dt \tag{5}$$

One very useful measure, the integral of time multiplied by absolute error (ITAE), has been proposed by Graham and Lathrop[10–12]

$$E=\int_0^\infty t|e(t)|\,dt \tag{6}$$

Inspection of equations 2 through 6 shows that they are functionals, since once the function of $e(t)$, such as e^2, e, $|e|$, etc., is specified, together with its range, a number E is obtained.

If an integral such as one of these is regarded as a satisfactory measure of error, a criterion or rule for choice of the best servo, as parameters are varied, is that this integral shall be a minimum. However, nothing has yet been said to indicate *how* the measure or measures are to be selected. A suitable measure should: (a) be responsive to the characteristics which a designer must optimize in a given servo, and (b) be susceptible to convenient computer mechanization.

In this paper some general measures of servo error will be examined and an attempt will be made to show how specific forms may be selected and used to suit the special needs of a servo designer in typical situations. As mentioned in the foregoing, only transient or step input functions will be considered. A point which should be emphasized is that the criteria to be discussed may be applied to nonlinear as well as linear servos.

A General Mathematical Expression for Error Measure

One general form of a metric, or measure of error, is the functional

$$E=\int_0^\infty F[e(t),\,t]\,dt \tag{7}$$

where F is some function of $e(t)$, and of t, the time measured from the initiation of the transient input $r(t)$. Equation 7 clearly includes, as special cases, equations 3–6. For example, F may be the function

$$F[e(t),\,t]=t|e(t)| \tag{8}$$

so that the general form reduces to the ITAE of equation 6. Similarly, expressions can be found for F corresponding to the forms of equations 3, 4, and 5. In this paper the expression in equation 7 will be considered as the general form of the

---

Paper 57-1033, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 7–11, 1957. Manuscript submitted June 7, 1957; made available for printing August 15, 1957.

W. C. SCHULTZ and V. C. RIDEOUT are both with the University of Wisconsin, Madison, Wis.

The research reported in this paper was supported in part by a National Science Foundation Grant.

The authors acknowledge the assistance of Prof. Preston C. Hammer of the University of Wisconsin Numerical Analysis Laboratory during the writing of this paper, and also the efforts of several graduate students of the Electrical Engineering Department in proofreading the manuscript.

* A typical functional, relative to a function $f(x)$ over the range $a<x<b$, is given by the integral $\int_a^b f(x)dx$.

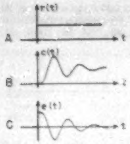


Fig. 1. Typical input, output, and error curves

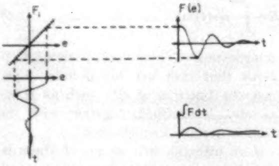


Fig. 2. Error measure for a criterion based on $F[e(t)] = e(t)$. $E = \int_0^\infty e(t)dt$

metric, and design criteria will be based on the minimization of particular forms of equation 7. Although the specific examples which have been cited to illustrate the general form have all resulted in *F*-functions which were easily expressed in a closed form, functions not so readily expressed will also be presented. Since a closed mathematical form will not then be convenient, a geometric form will be introduced. The idea of choosing a function, *F*, that will result in a measure which emphasizes certain features of the error $e(t)$ is suggested by Truxal[13] in his "importance-versus-error" curves, and by James, Nichols, and Phillips,[14] in their weight-of-error-versus-error curves.

Geometric Interpretation

The following form of the proposed criterion, which is somewhat less general than that given in equation 7, will be used to introduce the basic approach of the method. Let the metric be given by

$$E = \int_0^\infty F[e(t)]\ dt \qquad (9)$$

This form includes the first three integrals of the Introduction, the integrated error, the integral of squared error, and the integral of the absolute error, as special cases. Equation 9 defines a class of metrics within the general form and might be called the integral of a function of error (IFE).

The *F*-function can be interpreted geometrically, and, once this is done, the way will be opened for the endless possibilities of the method. To begin, consider the integrated error metric,

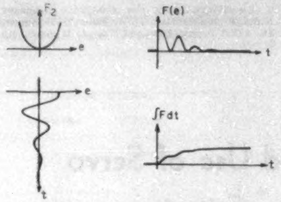


Fig. 3. Error measure for a criterion based on $F[e(t)] = e^2(t)$. $E = \int_0^\infty e^2(t)dt$

$$E = \int_0^\infty e(t)\ dt \qquad (10)$$

Fig. 1 shows a step input function, $r(t)$, and a typical output, $c(t)$. If the error is defined as in equation 1, it would appear as shown in Fig. 1(C). This error is next plotted on a composite figure, as illustrated in Fig. 2. The relationship between *F* and *e* is shown in the figure. This particular *F*-function, called F_1, simply translates $e(t)$ into itself, the integration process giving the area, *E*, under the $e(t)$ curve. Note that there is the possibility of a negative area for this metric. Now consider the metric given by equation 3,

$$E = \int_0^\infty e^2(t)\ dt \qquad (11)$$

the case illustrated in Fig. 3. The *F*-function here, called F_2, squares all ordinates of the error, $e(t)$; the integration gives the area under the squared-error curve. This is a form which is related to the well-known root-mean-square criterion. The third example, equation 5, is illustrated in Fig. 4. The function F_3 rectifies the error, and integration gives the area under the curve of absolute error. Although each geometric presentation shown has a simple corresponding analytical form, there is no reason to require only those geometric forms which do have simply expressible analytical forms. This fact gives much power to the proposed general method.

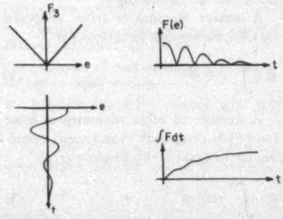


Fig. 4. Error measure for a criterion based on $F[e(t)] = |e(t)|$. $E = \int_0^\infty |e(t)|\ dt$

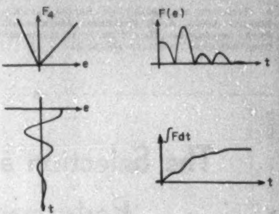


Fig. 5. Error measure for a criterion based on a modified absolute value of error

Suppose that the parameters of a servo are to be adjusted so that, for a step input, the absolute error will reduce to within some specified percentage as quickly as possible. The output of this system may show a steady oscillation, within the specified percentage limit, once this limit is first reached. The parameters of a given servo which are optimal for this case will be quite different from those for an "antiovershoot" or an "antiundershoot" case, to give two other possible examples. Another system that might be considered, the type-zero or regulator system, presents another difficulty. Since the type-zero system has a steady-state error for a step input, the integrals given by equations 3 through 6 would all fail to converge. These and other possible step-response requirements indicate the flexibility required in the general form of the metric or metrics upon which the criteria are to be based. The minimization of the integral of equation 9 is a criterion form which has the required flexibility, as will now be shown.

To illustrate the versatility of this criterion form, consider the antiovershoot case. Here the *F*-function of Fig. 4 might be modified as shown in Fig. 5 to penalize any overshoot. If overshoot is completely undesirable the *F*-function of Fig. 6 might be used, to put "full penalty" on overshoot error.

The first case mentioned, that of minimum time for the absolute error to decrease to within a specified percentage, would require the function shown in Fig. 7. In this function, typical of those with no time weighting, but in which time is to

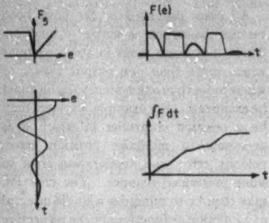


Fig. 6. Error measure for an antiovershoot criterion

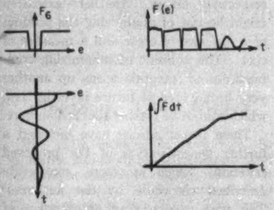


Fig. 7. Error measure for a criterion which permits a final error or error oscillation

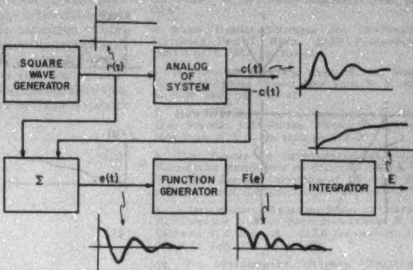


Fig. 8. Analog computer setup for determination of error measure, E. This setup can be used for optimization of the system

be minimized, all error lying beyond the specified limits is equally penalized. This criterion is also one that can be used in the study of type-zero system. (Some additional F-functions are discussed in the appendix.)

Computer Mechanization

The electronic differential analyzer is an invaluable tool for carrying out the computation required in the application of criteria based on equation 9. The type of computer used by the authors in this investigation is a computer of the high-speed or repetitive type, called the Wisconsin-Philbrick Computer. The computer was designed and constructed at the University of Wisconsin using Philbrick plug-in operational amplifiers and chopper-stabilizers. Because of the repetitive nature and the short (0.001-second) unit of time of the computer, the input device required is a square-wave pulse generator and the output device is a cathode-ray oscillograph. (All response curves shown in this paper are photographs of cathode-ray presentations.)

Refer to Fig. 8 for a diagram of the computer setup. The step-function is fed to the analog computer representation of the system being studied and also to a summing amplifier. The negative of the system output is added to the system input to produce the error, $e(t)$, which is then fed into a diode function generator to produce the desired $F[e(t)]$. The function generator output is then integrated and the integrator output is displayed on the oscillograph. The sequence of waveforms is illustrated in Fig. 8 for the metric of equation 5.

The method may now be summarized as follows:

1. Set the $F[e(t)]$ versus $e(t)$ characteristic on the function generator.
2. Apply $r(t)$ to the system and obtain $c(t)$.
3. Obtain $e(t) = r(t) - c(t)$.
4. Form $F[e(t)]$.
5. Integrate $F[e(t)]$ to obtain E.
6. Adjust the system parameters (within any given set of constraints) so that the value of E is minimized.
7. Read the set of values of parameters from computer.

From this summary it can be seen that the use of the method is relatively simple if an analog computer is available. It is particularly true if the computer can work repetitively at high speed.

Two things should be emphasized at

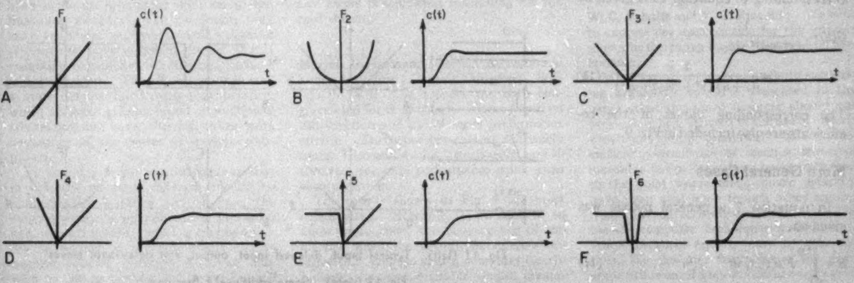


Fig. 9. Results of computer optimization of a fourth-order system for various criteria (equations 12 and 13)

A—$\alpha = 0.94$
$\delta = 0.1322$
$\omega_0 = 1.435$
B—$\alpha = 1.15$
$\delta = 0.324$
$\omega_0 = 1.480$

C—$\alpha = 0.94$
$\delta = 0.191$
$\omega_0 = 1.575$
D—$\alpha = 1$
$\delta = 0.16$
$\omega_0 = 1.687$

E—$\alpha = 1.325$
$\delta = 0.325$
$\omega_0 = 1.553$
F—$\alpha = 1.008$
$\delta = 0.205$
$\omega_0 = 1.595$

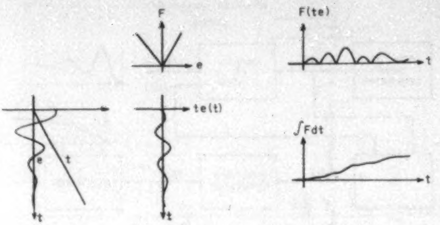


Fig. 10. A geometric interpretation of the ITAE

this point. First, the system being studied need not be linear, and second, the F-functions need not be linear. Since the electronic differential analyzer can deal with many kinds of nonlinearities, the servo designer has a powerful tool and a powerful method at his disposal. Nonlinear as well as linear systems can be studied and nonlinear or linear methods can be employed. Transients other than the step inputs considered here may, of course, be examined.

Results

The results of computer studies of the cases shown in Figs. 2 through 7 are tabulated in Fig. 9. Fig. 9 shows several geometric forms of the F-function along with the transient response resulting when the corresponding integral is minimized. For purpose of illustration, all cases presented are for a system with a transfer function given by

$$\frac{C(s)}{R(s)}=\frac{1}{s^4+a_1s^3+a_2s^2+a_3s+1} \qquad (12)$$

The open-loop transfer function $G(s)$ corresponding to equation 12 is given by

$$G(s)=\frac{1}{s(s^3+a_1s^2+a_2s+a_3)}$$
$$=\frac{1}{s(s+\alpha)\,(s^2+2\delta\omega_0s+\omega_0^2)} \qquad (13)$$

The corresponding values of the coefficients are also included in Fig. 9.

More General Cases

In equation 7 a general metric was given as

$$E=\int_0^\infty F[e(t),\,t]\;dt \qquad (14)$$

This form will now be discussed in more detail. The first example to be considered is the ITAE, which is formulated as

$$E=\int_0^\infty t|e(t)|\,dt \qquad (15)$$

This integral may be rewritten as

$$E=\int_0^\infty |te(t)|\,dt \qquad (16)$$

or, more generally,

$$E=\int_0^\infty F\{t\,[e(t)]\}\,dt \qquad (17)$$

This F-function can be interpreted as the same one used in Fig. 4 except that it is a function of $te(t)$ rather than of just $e(t)$ alone. In other words, the error input to the function generator is multiplied by the weighting function, t. The ITAE may therefore be represented geometrically as shown in Fig. 10. This representation now suggests a large class of metrics of the form $F\{t[e(t)]\}$, where the F-functions may take on the forms in Figs. 3 through 7, 9, or 12. The difference between this class and that described by equation 7 is then one of multiplying the error by t before applying to the input of the F-function generator. An even larger class of metrics of the form $F[w(t)e(t)]$ exists, where $w(t)$ is an arbitrary weighting function. Functions that might be used as weighting functions are: t^2, t^3, te^{-kt}, etc.

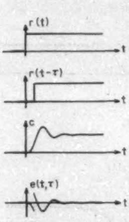
All cases discussed so far have been metrics based on a single F-function. There may arise cases in some systems where more than one output exists, or where more than one system variable must be examined. For example, a system may be a position controller in which it is necessary to minimize position error, velocity error, and acceleration error in some prescribed manner. The criterion may then be to minimize a single integral of a function or functions of the respective errors, or, possibly, to minimize a sum or product of integrals of functions of the respective errors. Another extension might be one of minimizing the product of a measure of error and a measure of cost. The concept of minimizing combinations of integrals opens up another very broad area of future investigation, which cannot be further touched on here.

These investigations have included a further generalization of the proposed criterion, which is more completely described elsewhere by the authors.[15] This work resulted from an extension of work carried out by Spooner and Rideout with the generalized error function.[3] Their GEF, for random, stationary inputs, is defined as

$$E(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[r(t-\tau)-c(t)]^2dt \qquad (18)$$

where $r(t)$ is the system input, $r(t-\tau)$ is the system input delayed by τ seconds and $c(t)$ is the system output. τ may be positive, negative, or zero. The extension to the transient input case may then be formulated as

$$E(\tau)=\int_{-\infty}^{\infty}F[e(t,\tau),\,t,\,\tau]dt \qquad (19)$$



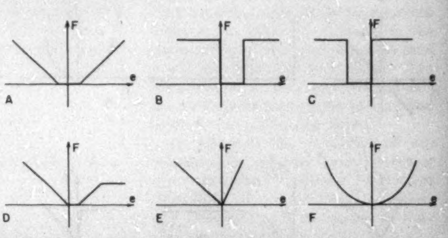


Fig. 11 (left). Typical input, delayed input, output, and delay-error curves

Fig. 12 (right). Some additional F-functions

A—Allows small final error, emphasizes all error greater than maximum allowable. The emphasis increases linearly as the magnitude of error increases
B—Antiovershoot. Allows small steady-state error, but not an overshoot
C—Antiundershoot. Allows small steady-state error, but not an undershoot
D—Ignores large initial error
E—Relative antiundershoot
F—Example of raising error e(t) to some power: $[e(t)]^{1.5}$

where $e(t,\tau)=r(t-\tau)-c(t)$. This definition of error, $e(t,\tau)$, called the delay-error, allows for a delay, τ, as shown in Fig. 11. The metric, $E(\tau)$, is therefore a function of the delay, τ, and the criterion may then require that τ as well as the system parameter values be varied to minimize the integral.

Special cases of equation 19 include Spooner's GEF for transient inputs,

$$E(\tau)=\int_{-\infty}^{\infty}[e(t,\tau)]^2dt \qquad (20)$$

and the ITAE of Graham and Lathrop,

$$E(O)=\int_{-\infty}^{\infty}t|e(t)|dt \qquad (21)$$

formed by setting τ equal to zero.

Appendix. Other Metrics

Fig. 12 is included to show how the geometric interpretation of measurement of error can be extended to include many other cases. A variation of the metric given in Fig. 7 is shown in Fig. 12(A). This metric places an emphasis on any error greater than some maximum allowable error. The functions of Fig. 12(B) and Fig. 12(C) are further modifications of Fig. 7. In Fig. 12(B) the same "amount" of error is allowed, but there is a strong discrimination against overshoot. In Fig. 12(C) there is a similarity to Fig. 12(B) except that "undershoot" is discriminated against. Fig. 12(D) is an attempt to set up a metric that pays little attention to the large initial error, since $F[e(t)]$ places a limit on maximum of $e(t)$. Many possibilities exist for the extension of this concept. Fig. 12(E) is related to Fig. 5 but places emphasis on the "undershoot" error. Figure 12(F) is included to show that there might be another broad class of metrics given by

$$E_n=\int_0^{\infty}F_n[e(t)]dt \qquad (22)$$

or

$$E_n=\int_0^{\infty}|e(t)|^n dt \qquad (23)$$

where n need not necessarily be an integer. For example, it might be of interest to obtain plots of E_n versus n, over a certain range of n, for a given system.

Many other F-functions might be added to those in Fig. 12 to include metrics based on other special system requirements.

References

1. **Methods of Mathematical Physics** (book), **R. Courant, D. Hilbert.** Interscience Publishers, Inc., New York, N. Y., vol. I, 1953, pp. 167–69.

2. **The Analysis and Synthesis of Linear Servomechanisms** (book), **A. C. Hall.** Technology Press, Massachusetts Institute of Technology, Cambridge, Mass., 1943, pp. 19–27.

3. **Correlation Studies of Linear and Nonlinear Systems, M. G. Spooner, V. C. Rideout.** *Proceedings,* National Electronics Conference, Chicago, Ill., vol. 12, 1956, pp 321–35.

4. **The Dynamics of Automatic Controls** (book), **R. C. Oldenbourg, H. Sartorious.** American Society of Mechanical Engineers, 1948, pp 66–67.

5. **Some Design Criteria for Automatic Controls, Paul T. Nims.** *AIEE Transactions,* vol. 70, pt. I., 1951, pp. 606–11.

6. **A Note on Control Area, T. M. Stout.** *Journal of Applied Physics,* New York, N. Y., vol. 21, Nov. 1950, pp. 1129–31.

7. **How to Specify the Performance of Closed-Loop Systems, J. E. Gibson.** *Control Engineering,* New York, N. Y., Sept. 1956, pp. 122–29.

8. **The Theory of Servomechanisms, With Particular Reference to Stabilization, A. L. Whiteley.** *Journal,* Institution of Electrical Engineers, London, England, vol. 93, 1946, pp. 353–72

9. **A Differential-Analyzer Study of Certain Nonlinearly Damped Servomechanisms, R. R. Caldwell, V. C. Rideout.** *AIEE Transactions,* vol. 72, pt. II, July 1953, pp. 165–69.

10. **The Synthesis of "Optimum" Transient Response: Criteria and Standard Forms, Dunstan Graham, R. C. Lathrop.** *Ibid.,* Nov., pp. 273–88.

11. **The Transient Performance of Servomechanisms with Derivative and Integral Control, Richard C. Lathrop, Dunstan Graham.** *Ibid.,* vol. 73, Mar. 1954, pp. 10–17.

12. **The Influence of Time Scale and Gain on Criteria for Servomechanism Performance, Dunstan Graham, Richard C. Lathrop.** *Ibid.,* July, pp. 153–58.

13. **Control System Synthesis** (book), **J. G. Truxal.** McGraw-Hill Book Company, Inc., New York, N. Y., 1955, pp. 413–14.

14. **Theory of Servomechanisms** (book), **H. James, N. B. Nichols, R. S. Phillips.** McGraw-Hill Book Company, Inc., 1947, pp. 308–11.

15. **A General Criterion For Servo Performance, W. C. Schultz, V. C. Rideout.** *Proceedings,* National Electronics Conference, vol. 13, 1957.

16. **An Integral Criterion for Optimizing Duplicator Systems on the Basis of Transient Response, Jack H. Crow.** *Doctor of Science Thesis,* Electrical Engineering Department, Washington University, St. Louis, Mo., June 1957.

◆

Discussion

S. P. Higgins (Minneapolis-Honeywell Regulator Company, Brown Instruments Division, Philadelphia, Pa.): The authors have made an interesting addition to the literature on criteria. In particular, they have considered nonsymmetrical weighting of error and have indicated the use of this weighting to penalize selectively overshoot and undershoot. They have also indicated a more definite break with past practice, which accepted criteria based on analytical advantages and have, thereby, taken more advantage of the power of modern computers.

Industrial process control requires that in different instances different criteria be applied. For instance, in one case, the emphasis may be on minimizing overshoot and, in another case, emphasis may be on reducing cycling to prevent effects due to coupled control systems. Also, we may wish to determine how well a system will satisfy two criteria separately by adjustment of instrument parameters. For these purposes, the approach described by the authors is of practical value.

It is, perhaps, well to keep in mind that, while simple criteria and results derived from them are needed in practical design work, the criteria on which performance is ultimately judged are generally more complex. Sometimes they can be expressed in quantitative terms. This subject has, unfortunately, not been well developed in the literature.

A minor point is brought up by Fig. 6, where the limiting on the F_5 values for negative errors is somewhat misleading for the case shown.

Morton G. Spooner (Cornell Aeronautical Laboratory, Inc., Cornell University, Buffalo, N. Y.): The authors are to be congratulated for their clearly written paper on the selection and use of servo performance criteria. The paper presents an extremely useful procedure for evaluating servo transient response performance using some chosen criteria.

The results shown in Fig. 9 are most enlightening. However, I would like to know more about the convergence of the metrics when the system parameters are adjusted. Did the analog computer clearly indicate the most desirable system parameters in all the cases considered?

The authors state that the system under test need not be linear, and also that the F-functions need not be linear. An interesting study in the nonlinear case would be one in which the input signal was contaminated by a random noise with a zero mean and a Gaussian distribution. The error in this case would be the uncontaminated input minus the output. Do the authors believe that the addition of noise to the input would change the optimum system parameters when some specified metric was used?

W. C. Schultz and **V. C. Rideout:** We wish to express our appreciation for the interest shown in the paper by Mr. Higgins and Mr. Spooner.

Mr. Higgins' remarks contribute supporting arguments to those presented in the paper, that the servo designer should not necessarily be restricted to the use of criteria which are simply analytical. The mathematical convenience of mean-square-error measures has perhaps been overemphasized to the point where mean-square measures have been used even though other metrics would have given more useful criteria. The use of computer techniques makes many other measures equally easy to use, and thus, the designer may choose the most appropriate one if such devices are employed.

In regard to Fig. 6, it is agreed that it is perhaps misleading. For the particular error curve shown, the F_5 curve is satisfactory, since the error does not exceed the "range" of F_5. If, however, the error for some other system was such that the error exceeded the maximum value shown in Fig. 6, a limit would also have to be placed on the portion of the *F*-function in the right-hand quadrant. This limit would most likely be at the same level as that shown in the left-hand quadrant.

In regard to the remark concerning Fig. 9, we wish to thank Mr. Spooner for raising this important point. For the particular system presented, the fourth-order linear system, the computer did clearly indicate that the responses shown in Fig. 9 are the "best" system responses, relative to the several metrics. However, we wish to point out that for higher-order systems and the measure of Fig. 9(C), a series of "relative minima" were found to exist. Specifically, for an eighth-order system, the computer very quickly showed *a* minimum, but other minima were also found, and all of these minima had to be examined to establish the absolute minimum. Future work in this area will include a more rigorous mathematical approach to the problem of uniqueness.

Certain difficulties are encountered if the case where the possibility of contamination of the input signal by noise is considered. The immediate difficulty is that a continuous noise signal present at the output will result in infinite integrals. One possible modification is the use of a finite limit of integration time. Another approach would be the use of an *F*-function of the type shown in Fig 9(F) or Fig. 12(A). Although in a theoretical sense the integral would still become infinite, due to large peak noise signals, in a practical case the peak amplitudes would be limited at the output, so that an *F*-function of the type mentioned would satisfactorily handle the cases where noise was present. In general it would be expected that the "optimum" system parameters would be different for noise cases, for any given error measure.

Extended Switching Criterion for Second-Order Saturated Servomechanisms

J. W. DIESEL
STUDENT MEMBER AIEE

A PRIMARY requirement in many positioning systems is that errors resulting from large input disturbances be reduced to tolerable values as quickly as possible. For a linear system, this requirement could be satisfied by making the ratio of torque to inertia sufficiently large. However, even with the best servomotor design this ratio is limited; therefore, to sufficiently reduce the response time, the designer must frequently contend with saturation in the system components and resulting saturated values of available torque. This paper deals with this saturated type of operation.

If saturation effects are present in a given control system, it is reasonable to expect that optimum response in the sense of minimum recovery time would be obtained if the corrective torque is allowed to saturate at its maximum permissible value during recovery from the disturbance, so that full available power is utilized at all times. To obtain optimum response it is also necessary that this torque be switched to its full reverse value at the proper instant during recovery. The proper instant of torque reversal is determined by a computer, which utilizes information obtained from the input and output, and manipulates the corrective torque accordingly.

Computer characteristics have been devised[1-7] for systems employing a constant maximum torque and subjected to step-displacement inputs only. Most actual systems, however, are incapable of maintaining a constant torque since the servomotor torque is usually a function of the output speed; also, most systems are excited by more complex inputs than simple displacement steps.

Accordingly, this paper extends present knowledge in two directions:

1. Controller characteristics are established for inputs composed of a combination of velocity and displacement steps.

2. Controller characteristics are established for servomotors with actual speed-torque characteristics.

This study is restricted to plants describable by a second-order equation of motion. For most positioning systems this assumption is not unreasonable. Another limitation is that the equations contain no terms dependent on the output displacement. This is necessary in order that the controller characteristics will not depend on the magnitude of the input, but only on the error between the input and the output. This requirement is equivalent to assuming no spring forces, which is usually the case for systems subjected to velocity inputs. The conclusions of the investigation are therefore quite generally applicable.

Nomenclature

c = normalized output or controlled variable
r = normalized input or reference variable
e = normalized error, $r-c$
m = normalized control voltage or manipulated variable
τ = normalized time
$\Phi(m,c')$ = normalized torque as function of m and c'
$_0$ = (subscript) indicates initial value of a variable
max = (subscript) indicates maximum value of a variable
$'$ = (superscript) indicates first derivative with respect to normalized time
$''$ = (superscript) indicates second derivative with respect to normalized time
$f(e,e',r')$ = function of e,e',r' defined in text
$g(u)$, $h(u)$ = functions of u defined in text
δ = function of e' defined in text
v = normalized speed and dummy variable of integration
θ_0 = output displacement, radians
θ_i = input displacement, radians
V = control voltage, volts
t = time, seconds
T = developed torque, foot-pounds
$\dot{}$ = (superscript) indicates first derivative with respect to time, seconds
$\ddot{}$ = (superscript) indicates second derivative with respect to time, seconds
K_m = motor gain constant, foot-pounds per volt
ζ_m = motor damping constant, foot-pound-seconds
J = inertia, foot-pound-seconds squared
ζ_r = viscous damping coefficient, foot-pound-seconds
ζ_c = coulomb friction coefficient, foot-pound
ρ_c = normalized coulomb friction ratio
n = rotor speed, revolutions per minute
n_s = synchronous speed, revolutions per minute
R_1 = stator resistance per phase, ohms
X_1 = stator reactance per phase at line frequency, ohms
R_{2e} = equivalent rotor resistance per phase referred to stator, ohms
X_{2e} = equivalent rotor reactance per phase at line frequency, ohms
V_1 = main winding voltage, volts

System Studied

As stated previously, the investigation is directed to systems excited by a special class of inputs. Explicitly, it is assumed that the input displacement and velocity

---

Paper 57-1036, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 7-11, 1957. Manuscript submitted June 10, 1957; made available for printing August 22, 1957.

J. W. Diesel is at Washington University, St. Louis, Mo.

The material presented in this paper is based on the thesis submitted by the author in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering at Washington University. Acknowledgment is due Professor John Zaborszky for supervising the thesis and reviewing the paper.

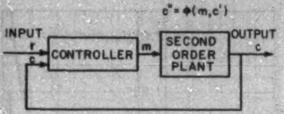


Fig. 1. Schematic diagram of general second-order positioning system

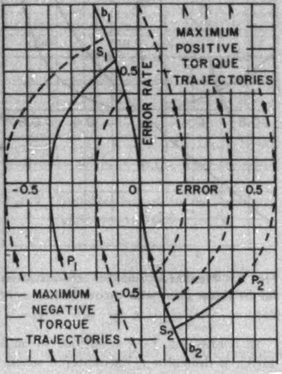


Fig. 2. Maximum torque trajectories in the phase plane

are continuous except at step-type disturbances, and that the system is allowed sufficient time to recover between such disturbances. Further, it is assumed that the input velocity does not change appreciably during the recovery period, so that only inputs with constant velocity need be considered. It then suffices to consider a single input of the form

$$r=r_0+r_0'\tau, \quad 0<\tau \tag{1}$$

with the initial conditions

$$c=0,\ c'=c_0' \text{ at } \tau=0 \tag{2}$$

The optimum control to be considered applies to any second-order plant for which the output acceleration c'' is a single-valued function of both the manipulated variable m (see Fig. 1) and the output speed c', but not the output displacement

$$c''=\phi(m, c') \tag{3}$$

This definition includes the broad class of nonlinear characteristics of such servomotors as are used in control systems of the standard type. To simplify the analysis, it is convenient to impose the additional restrictions

$$\phi(-m, -c')=-\phi(m, c') \tag{4}$$

$$\phi(m, c')<\phi(m_{\max}, c'),\ m<m_{\max} \tag{5}$$

$$\phi(m_{\max}, c')>0,\ c'<c_{\max}' \tag{6}$$

$$\phi(m_{\max}, c')<0,\ c_{\max}'<c' \tag{7}$$

The symmetry condition 4 ensures that the motor can accelerate equally well in either direction. Condition 5 states that the maximum acceleration at any speed occurs at a specified maximum value of the manipulated variable m, which is frequently a control voltage. The maximum speed attainable by the system is defined by conditions 6 and 7. Also, the variables are assumed normalized, so that

$$m_{\max}=1,\ c_{\max}'=1,\ \phi(m_{\max}, 0)=1$$

Optimization Problem

If the input disturbance defined by equations 1 and 2 is applied to the plant, equations 3 through 7, an error in position results at $\tau=0$. Since a second-order system does not overshoot when the error and error rate are zero at the same instant, optimum control may then be defined as that method of control which simultaneously reduces the foregoing error and error rate to zero in the least possible time.

The optimum method of control is best understood by referring to the phase plane of Fig. 2, in which the co-ordinates are the error and error rate existing in the system following the disturbance.

In general, the corrective torque is controlled by manipulating the variable m, the resultant response of the system then being represented by a corresponding trajectory in the phase plane. A family of such curves is obtained by setting m at its maximum positive value $m_{\max}$ and prescribing various initial conditions. These curves are referred to as maximum positive torque trajectories. Similarly, a family of negative torque trajectories is obtained by setting $m=-m_{\max}$. One of the maximum positive torque trajectories passes through the origin. The portion of this curve in the upper-half plane represents motion toward the origin and is labeled b_1 in Fig. 2. Similarly, the portion of the negative maximum torque trajectory terminating at the origin and in the lower-half plane is labeled b_2. Together these curves constitute a boundary dividing the plane into two regions. This boundary is called the switching boundary for reasons which will immediately be apparent.

The optimum method of control may now be described as follows:

> Whenever the error and error rate of the system are such that the representative point of the system lies to the left of the switching boundary in the phase plane, $-m_{\max}$ should be applied. Whenever the representative point lies to the right of this boundary, $+m_{\max}$ should be applied.

In Appendix I it is shown that such manipulation of m indeed results in the desired optimum control. It follows that whenever the point representing the initial error and error rate lies to the left of the switching boundary, the representative point subsequently moves along a maximum negative torque trajectory with $m=-m_{\max}$ until b_1 is crossed. The variable m is then switched to $+m_{\max}$ and the final stage of the trajectory is along b_1 into the origin, where the error and error rate are simultaneously zero, and the system is recovered from the disturbance. The complete trajectory during recovery would then be represented by a curve such as P_1S_1O in Fig. 2. The correspond-

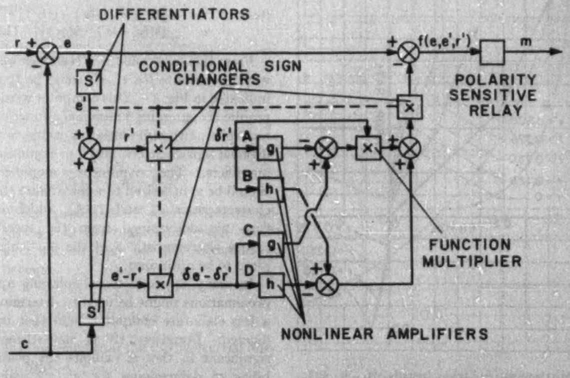


Fig. 3. Schematic diagram of computer required for optimum control

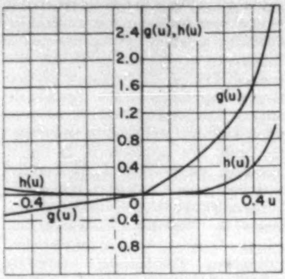


Fig. 4. Nonlinear amplifier characteristics for plant with coulomb and viscous friction

ing trajectory when the initial point lies to the right of the switching boundary is obvious, with the final portion of the trajectory along b_2. Such a trajectory is exemplified by curve P_2S_2O of Fig. 2.

The equation of the switching boundary is derived in Appendix II and is given by

$$e=\delta\int_{\delta r'}^{\delta r'-\delta e'}\frac{\delta r'-v}{\phi(+m_{\max},v)}dv,\ \delta=\begin{cases}+1,\ e'>0\\-1,\ e'<0\end{cases}\tag{8}$$

Note that the switching boundary depends on r' the velocity of the input, and in this respect the switching criterion differs from that for systems disturbed by step displacements only, the switching boundary being independent of the input in the latter case.

The point with co-ordinates e, e' lies to the right or left of the switching boundary depending on whether the quantity

$$f(e,e',r')=e-\delta\int_{\delta r'}^{\delta r'-\delta e'}\frac{\delta r'-v}{\phi(+m_{\max},v)}dv\tag{9}$$

is positive or negative, respectively. Accordingly, the desired manipulation of the corrective torque may be obtained if the computed quantity $f(e, e', r')$ is fed to a polarity-sensitive relay with output $\pm m_{\max}$ determined by the algebraic sign of $f(e, e', r')$. The resulting system would then provide optimum response to step displacement and velocity inputs.

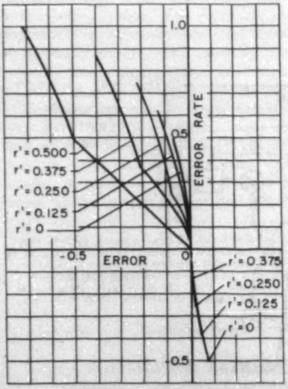


Fig. 5. Switching boundaries for a linear plant with viscous and coulomb friction

Although the expression for $f(e, e', r')$ contained in equation 9 is exceedingly compact, it offers no suggestion as to how this quantity would be computed in an actual system. Equation 9 is therefore rearranged by dividing the range of integration to yield

$$f(e,e',r')=e-\delta\left\{\delta r'\left[\int_0^{\delta r'-\delta e'}\frac{dv}{\phi(m_{\max},v)}-\int_0^{\delta r'}\frac{dv}{\phi(m_{\max},v)}\right]-\left[\int_0^{\delta r'-\delta e'}\frac{vdv}{\phi(m_{\max},v)}-\int_0^{\delta r'}\frac{vdv}{\phi(m_{\max},v)}\right]\right\}\tag{10}$$

By defining

$$g(u)=\int_0^u\frac{dv}{\phi(m_{\max},v)}\tag{11}$$

and

$$h(u)=\int_0^u\frac{vdv}{\phi(m_{\max},v)}\tag{11A}$$

$f(e, e', r')$ may finally be written

$$f(e,e',r')=e-\delta\{\delta r'[g(\delta r'-\delta e')-g(\delta r')]-[h(\delta r'-\delta e')-h(\delta r')]\}\tag{12}$$

From this expression it is seen that one way to compute $f(e, e', r')$ could be that indicated in Fig. 3. This computer would require six summing junctions, a function multiplier, two differentiators, three conditional sign changers, and four nonlinear amplifiers. The nonlinear amplifiers should be synthesized to approximate the characteristics 11 and 11(A), which in turn are determined from the speed-torque characteristic 3 of the particular plant to be optimized.

In many applications, simplifying approximations might be used to determine a less elaborate computer than that indicated. Equation 12 is nevertheless significant in that it exhibits the possibility of determining $f(e, e', r')$ using only existing computer techniques.

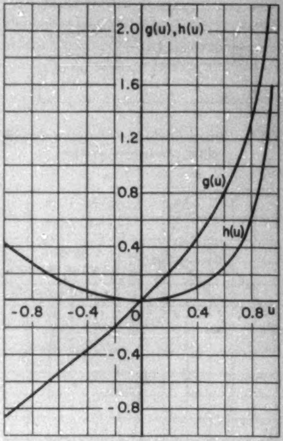


Fig. 6. Nonlinear amplifier characteristics for the linear plant with viscous friction only

In connection with equations 11 and 11(A), a difficulty is encountered whenever the argument u approaches unity, since both $g(u)$ and $h(u)$ become unbounded in this case. Physically, this means that two of the amplifiers indicated in Fig. 3 would saturate if the input velocity were to approach or exceed the maximum output speed of the system. In this event the proper sense of corrective torque is assured if the saturated output of amplifier A in Fig. 3 is made sufficiently larger than that of amplifier B.

Linear Plant with Viscous and Coulomb Friction

Consider a motor described by the linear speed-torque characteristics

$$T=K_mV-\zeta_m\dot{\theta}_0,\ |V|\leqq V_{\max}\tag{13}$$

where the load consists of viscous and coulomb friction as well as inertia

$$T=J\ddot{\theta}_0+\zeta_r\dot{\theta}_0+\zeta_c\frac{\dot{\theta}_0}{|\dot{\theta}_0|}\tag{14}$$

The variables may be normalized by setting

$$c=\frac{(\zeta_m+\zeta_r)^2}{K_mV_{\max}J}\theta_0$$

$$m=V/V_{\max}$$

$$r=\frac{(\zeta_m+\zeta_r)^2}{K_mV_{\max}J}\theta_i$$

$$\tau=\frac{(\zeta_m+\zeta_r)}{J}t$$

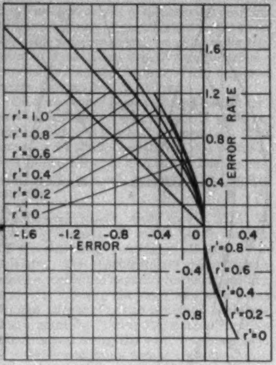


Fig. 7. Switching boundaries for linear plant with viscous friction only

so that equations 13 and 14 become

$$c''=m-c'-\rho_c\frac{c'}{|c'|},\ \rho_c=\frac{\zeta_c}{K_m V_{\max}} \tag{15}$$

By comparing equation 15 with equation 3

$$\phi(m,\,c')=m-c'-\rho_c\frac{c'}{|c'|}$$

the nonlinear amplifier characteristics are then given by

$$g(u)=\int_0^u\frac{1}{1-v-\rho_c\dfrac{v}{|v|}}\,dv$$

$$=-\ln\left(1-\frac{u}{1-\rho_c\dfrac{u}{|u|}}\right)$$

$$h(u)=\int_0^u\frac{v}{1-v-\rho_c\dfrac{v}{|v|}}\,dv=-u-$$

$$\left(1-\rho_c\frac{u}{|u|}\right)\ln\left(1-\frac{u}{1-\rho_c\dfrac{u}{|u|}}\right)$$

These characteristics are shown in Fig. 4 for $\rho_c=0.5$. The equation of the switching boundary is obtained by substituting the foregoing expressions into equation 12, and setting the result equal to zero.

$$e=\left\{\delta\left[1-\rho_c\frac{\delta(r'-e')}{|\delta(r'-e')|}\right]-r'\right\}\times$$

$$\ln\left[\frac{1-\rho_c\dfrac{\delta(r'-e')}{|\delta(r'-e')|}-\delta(r'-e')}{1-\rho_c\dfrac{\delta(r'-e')}{|\delta(r'-e')|}}\right]-$$

$$\left\{\delta\left[1-\rho_c\frac{\delta r'}{|\delta r'|}\right]-r'\right\}\times$$

$$\ln\left[\frac{1-\rho_c\dfrac{\delta r'}{|\delta r'|}-\delta r'}{1-\rho_c\dfrac{\delta r'}{|\delta r'|}}\right]-e'$$

These boundaries are shown in Fig. 5 for positive input velocities. The switching boundaries for negative input velocities are similar.

An important special case occurs when $\rho_c=0$, corresponding to a linear plant with viscous friction only. The functions $g(u)$ and $h(u)$ are then given by

$$g(u)=-\ln(1-u)$$
$$h(u)=-\ln(1-u)-u$$

and the switching boundaries become

$$e=-e'+(\delta-r')\ \ln\left(1+\frac{e'}{\delta-r'}\right)$$

The amplifier characteristics and switching boundaries are shown in Figs. 6 and 7, which may be compared with Figs. 4 and 5 to observe the effect of coulomb friction.

Two-Phase Induction Motor

Because of the pronounced nonlinearity of its speed-torque characteristics and its frequent service in positioning systems, the 2-phase induction motor provides an especially practical illustration of the foregoing principles.

Before equations 11 and 12 can be applied, it is necessary to determine the speed-torque characteristics of the 2-phase induction motor, corresponding to equation 3. If the full control-winding voltage equals the main winding voltage, the torque may be expressed as

$$T=K\frac{1-v}{X^2v^2-2(X^2+Y^2)v+1} \tag{16}$$

where

$$K=\frac{33{,}000}{2\pi n_s\times746}\ \frac{2V_1^2R_{2e}}{[(X_1+X_{2e})^2+R_1^2]}$$

$$v=n/n_s$$

$$X^2=\frac{(X_1+X_{2e})^2+R_1^2}{(X_1+X_{2e})^2+(R_1+R_{2e})^2}$$

$$Y^2=\frac{R_1R_{2e}}{(X_1+X_{2e})^2+(R_1+R_{2e})^2}$$

If the quantities X^2 and Y^2 are so adjusted that the maximum torque occurs at a slip of two, so that equation 16 becomes

$$T=K\frac{1-v}{X^2v^2+(3X^2-1)v+1}$$

It is assumed that the load is inertial only, so that

$$T=J\ddot{\theta}_0 \tag{17}$$

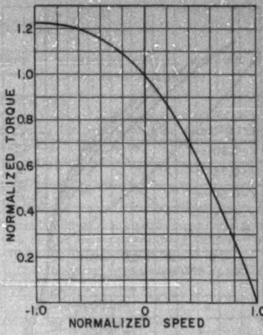


Fig. 8. Speed-torque characteristic of 2-phase servomotor with full control-winding voltage

The variables are next normalized by defining

normalized time $\tau=\dfrac{Kn_s}{J2\pi\times60}\,t$

normalized output $c=\dfrac{Kn_s^2}{J(2\pi\times60)^2}\,\theta_0$

Equations 16 and 17 then combine to give

$$c''=\phi(m_{\max},\,c')=\frac{1-c'}{X^2c'^2+(3X^2-1)c'+1}$$

The nonlinear amplifier characteristics

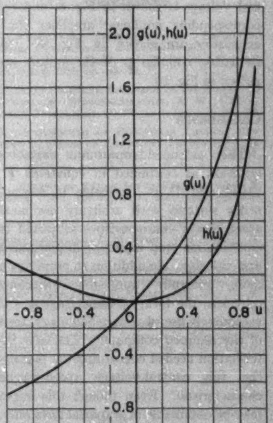


Fig. 9. Nonlinear amplifier characteristics of computer for 2-phase servomotor

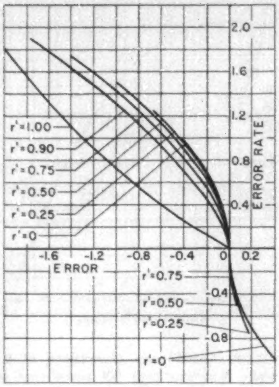


Fig. 10. Switching boundaries for the 2-phase servomotor

are obtained from equations 11 and 11(A)

$$g(u) = \frac{-X^2u^2}{2} - (4X^2-1)u - 4X^2 \ln(1-u)$$

$$h(u) = \frac{-X^2u^3}{3} - \frac{(4X^2-1)}{2}u^2 - 4X^2u - 4X^2 \ln(1-u)$$

If $R_1=2$ ohms, $R_{2e}=4.02$ ohms, $X_1=1$ ohm, $X_{2e}=1$ ohm, then $X^2=0.18525$. For these values of the parameters the speed-torque characteristic with full control voltage is shown in Fig. 8, and the corresponding nonlinear amplifier characteristics are shown in Fig. 9. The switching boundaries of the system are shown in Fig. 10.

Conclusions

The extended optimum switching criterion as contained in equations 11, 11(A) and 12 is applicable to second-order systems with arbitrary nonlinear speed-torque characteristics subjected to combination step inputs of displacement and velocity. In addition to providing a more generally applicable switching criterion, the form of these equations is exceedingly useful as a guide for synthesis of the required computer. The only information needed concerning the plant is the full control-voltage speed-torque characteristic. Even though this curve may be obtainable in graphical form only, the functions $g(u)$ and $h(u)$ may be plotted therefrom, and the required nonlinear amplifiers synthesized using only these plots as a guide.

Although the complexity of the resulting computer would therefore appear to be relatively independent of the particular plant to be optimized the instrumentation involved would not be justified in many applications. The foregoing theory nevertheless might serve as a guide in designing less than optimum systems.

It should finally be observed that the saturated type of operation described in this paper is only a means of reducing the error to relatively small values following large input disturbances. Once the error is small, the system should revert to some type of continuous control. This second mode of control is necessary to provide satisfactory response to small disturbances, and to reduce errors incurred during the saturated mode of operation due to errors in instrumentation and small changes in input velocity between major disturbances.

Appendix I. Determination of Optimum Method of Control

Consider the response of the system shown in Fig 1 following an input disturbance with constant velocity r'. It will be shown that if m is manipulated according to the criterion described in the text, the resulting response is optimum in the sense of minimum recovery time. The initial conditions corresponding to the disturbance determine the initial co-ordinates e and e' of the representative point in the phase plane. This point is denoted by P_0. The trajectory in the phase plane must extent from this point to the origin, if the system is to finally recover from the disturbance, and the exact shape of the trajectory between these points is to be determined by the manner in which m is varied during recovery.

If P_i and P_{i+1} represent points along such a trajectory, the time required to move along the curve from P_i to P_{i+1} is denoted by $\tau(P_i, P_i+_1)$. If the complete trajectory from P_0 to the origin is divided into sections by the points $P_1, P_2, \ldots, P_n$, the total time required to traverse the trajectory is $\tau=\tau(P_0, P_1)+\tau(P_1, P_2)+\ldots+\tau(P_n,0)$. The problem is then to determine the trajectory for which τ is a minimum.

Several properties of the phase-plane trajectories are now noted:

1. The time required to move along a trajectory from a given point with abscissa e_1 to another point with abisscsa e_2 is a minimum if, along the trajectory, $|e'|$ is a maximum. This may be seen by writing

$$\tau_2-\tau_1=\int_{\tau_1}^{\tau_2} d\tau = \int_{e_1}^{e_2} \frac{1}{e'}\, de$$

where τ_1 is the time at which the error is e_1, and τ_2 is the time at which the error is e_2. If e' is positive, e increases with time and de is positive. Conversely, if e' is negative, de is also negative. The conclusion follows immediately.

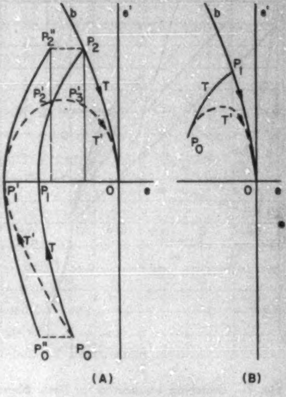


Fig. 11. Phase-plane trajectories corresponding to optimum and less than optimum control

2. There is a unique positive torque trajectory and a unique negative torque trajectory passing through each point of the phase plane for which $-1+r'<e'<1+r'$.

For a given fixed value of r' such that $|r'|<1$, the equation of motion of the system can be written as a second-order differential equation in e. The foregoing statement then follows from the theory of second-order differential equations, the limits on e' being a result of equations 6 and 7 which define the maximum output speed of the system.

3. The maximum torque trajectories for a given sense of torque all have the same shape except for a translation along the error axis. The maximum torque trajectories represent solutions of a differential equation in which there are no terms dependent on the displacement of the system. For a given fixed value of m, all such trajectories must then have the same shape except for their position along the error axis, as determined by initial conditions.

4. If the slope of a maximum torque trajectory through a point in the phase plane is positive, it exceeds the slope of any other trajectory through the point. Similarly, if the slope of the maximum torque trajectory is negative, any other trajectory through the point has a larger slope. Since r' is assumed constant, $r''=0$. Then

$$\phi(m, c') = c'' = r'' - e'' = -e''$$
$$= -\frac{de'}{de}\frac{de}{d\tau} = -\frac{de'}{de}e'$$

so that

$$\frac{de'}{de} = \frac{-\phi(m, c')}{e'}$$

Here de'/de is the slope of the trajectory through the point. If the ordinate of the point is e', then $c'=r'-e'$ is fixed, so that by equation 5, $\phi(m,c')$ is a positive maximum if $m=+1$, and a negative maximum if $m=-1$, from which the conclusion is readily deduced.

5. The slope de'/de of a maximum positive torque trajectory is negative above the error axis and positive below the axis, while the slope of a maximum negative torque trajectory is positive above the axis and negative below the axis.

On a maximum positive torque trajectory, $m=+1$ so that $c''=\phi(+1, c')$ is positive, and $(de'/de)e'=e''=-c''$ is negative. The slope de'/de is then negative if e' is positive, and positive if e' is negative. The conclusion when $m=-1$ follows in a similar manner.

Reference is now made to Fig. 11(A) in which it is assumed that P_0 lies to the left of the switching boundary. If $m=-1$ is initiated, the representative point moves along a maximum negative torque trajectory from P_0, crossing the error axis at a point P_1 and intersecting the switching boundary b at a point P_2. From P_2 the representative point moves along the switching boundary, a positive maximum torque trajectory, into the origin. The complete trajectory described is denoted by T. That the general shape of this curve is as shown follows from property 5. The time τ required for recovery of the system in this case is

$$\tau=\tau(P_0, P_1)+\tau(P_1, P_2)+\tau(P_2, 0)$$

Since T represents the method of control outlined in the text, it must be shown that the time required to traverse any other trajectory T' from P_0 to 0 will exceed τ.

If $m>-1$ along some other trajectory T' from P_0, properties 4 and 5 ensure that the trajectory crosses the e axis at a point P_1' to the left of P_1. The trajectory T' must then continue from P_1' to 0 as indicated. Using property 2, the maximum negative torque trajectory through P_1' may be obtained by shifting $P_0P_1P_2$ to the left. This curve is shown as $P_0''P_1'P_2''$. From properties 4 and 5 it is clear that T' lies above $P_0''P_1'$ and below both $P_1'P_2''$ and P_2O, as indicated. Perpendiculars from P_2'' and P_2 to the error axis therefore cross T' at points P_2' and P_3' as shown.

Using properties 1 and 3 gives

$$\tau(P_0, P_1')>\tau(P_0'', P_1')=\tau(P_0, P_1)$$

$$\tau(P_1', P_2')>\tau(P_1', P_2'')=(P_1, P_2)$$

$$\tau(P_2', P_3')>0$$

$$\tau(P_3', 0)>\tau(P_2, 0)$$

Since the time τ' required to traverse T' is

$$\tau'=\tau(P_0, P_1')+\tau(P_1', P_2')+\tau(P_2', P_3')+\tau(P_3', 0)$$

it is seen by adding the foregoing inequalities that $\tau'>\tau$, which is the desired result.

If P_0 lies above the e axis, as in Fig. 11(B), properties 1 through 5 ensure that the maximum torque trajectory T lies above any other trajectory from P_0 to 0, and that the recovery time along T is therefore a minimum.

The case where the initial point P_0 lies to the right of the switching boundary is similar to the foregoing cases.

Appendix II. Derivation of Equation of the Switching Boundary

It is desired to determine the relation which exists between e and e' along a maximum torque trajectory terminating at the origin of the phase plane. To this end, it may be assumed that the error and error rate are reduced to zero at time $\tau=0$, and that prior to this time the output quantity c satisfies the equation of motion 3 with $m=\pm1$. Corresponding values of $e=r-c$ and $e'=r'-c'$ must then be found for $\tau<0$.

Since there are no spring forces in the system, the boundary condition at the origin may be expressed as

$$c=r;\ c'=r',\ \text{at}\ \tau=0$$

with no loss of generality. The change of variables

$$v=c'$$

reduces equation 3 to the first-order equation

$$v\frac{dv}{dc}=\phi(m, v),\quad m=\pm1$$

Separating variables and integrating gives

$$c=\int_{r'}^{c'}\frac{v\,dv}{\phi(m, v)}$$

wherein the boundary conditions were used to obtain the limits of integration.

Again using the boundary conditions to establish limits of integration, τ may be obtained as

$$\tau=\int_0^\tau d\tau=\int_{r'}^{c'}\frac{1}{\frac{dc}{d\tau}\frac{dv}{dc}}\,dv$$

But

$$\frac{dc}{d\tau}\frac{dv}{dc}=v\frac{dv}{dc}=c''=\phi(m, v)$$

so that

$$\tau=\int_{r'}^{c'}\frac{1}{\phi(m, v)}\,dv$$

Since the integrand has the same algebraic sign as m according to equations 4 and 6, the foregoing expression represents negative values of time if

$$m=+1,\ \text{when}\ r'-c'>0$$

$$m=-1,\ \text{when}\ r'-c'<0$$

i.e.,

$$m=\delta=\begin{cases}+1, & e'>0\\ -1, & e'<0\end{cases}$$

Since r' is assumed constant and $r=0$ at $\tau=0$, the input must be given by

$$r=r'\tau,\quad \tau<0$$

or, using the expression for τ found in the foregoing, it is given by

$$r=r'\int_{r'}^{c'}\frac{1}{\phi(\delta, v)}\,dv,$$

where the limits of integration and definition of δ ensure that the resulting values of r correspond to negative values of τ.

The foregoing expressions for r and c may be used to obtain e.

$$e=r-c=r'\int_{r'}^{c'}\frac{1}{\phi(\delta, v)}\,dv-\int_{r'}^{c'}\frac{v\,dv}{\phi(\delta, v)}=\int_{r'}^{r'-e'}\frac{r'-v}{\phi(\delta, v)}\,dv$$

To eliminate δ from the integrand, the variable of integration is changed to $u=(1/\delta)v$. Since $\delta=\pm1$ is constant along the trajectory, $dv=\delta du$, so that

$$e=\int_{\delta r'}^{\delta r'-\delta e'}\frac{r'-\delta u}{\phi(\delta, \delta u)}\,\delta du$$

But $\phi(\delta, \delta u)=\delta\phi(1, u)$ from the symmetry condition 4. Also δ may be factored from the numerator of the integrand, using $\delta^2=+1$. The final result is

$$e=\delta\int_{r'}^{\delta r'-\delta e'}\frac{\delta r'-u}{\phi(1, u)}\,du$$

which is the desired relation between e and e'.

References

1. NONLINEAR TECHNIQUES FOR IMPROVING SERVO PERFORMANCE, D. C. McDonald. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 400–21.

2. A PHASE-PLANE APPROACH TO THE DESIGN OF SATURATING SERVOMECHANISMS, A. M. Hopkin. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

3. OPTIMIZATION OF NONLINEAR CONTROL SYSTEMS BY MEANS OF NONLINEAR FEEDBACKS, R. S. Neiswander, R. H. MacNeal. *Ibid.*, vol 72, pt. II, Sept. 1953, pp. 262–72.

4. ERRORS IN RELAY SERVO SYSTEMS, Louis F. Kazda. *Ibid.*, vol. 72, pt. II, Nov. 1953, pp. 323–28.

5. EFFECTS OF FRICTION IN AN OPTIMUM RELAY SERVOMECHANISM, T. M. Stout. *Ibid.*, vol. 72, pt. II, Nov. 1953, pp. 329–36.

6. AN INVESTIGATION OF THE SWITCHING CRITERIA FOR HIGHER ORDER CONTACTOR SERVOMECHANISMS, Irving Bogner, Louis F. Kazda. *Ibid.*, vol. 73, pt. II, July 1954, pp. 118–27.

7. ANALYSIS AND DESIGN PRINCIPLES OF SECOND AND HIGHER ORDER SATURATING SERVOMECHANISMS, R. E. Kalman. *Ibid.*, vol. 74, pt. II, Nov. 1955, pp. 294–310.

A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves

J. ZABORSZKY
MEMBER AIEE

H. J. HARRINGTON
ASSOCIATE MEMBER AIEE

AN EARLIER PAPER[1] introduced a method for handling multiple nonlinearities through describing functions, and developed such describing functions for single-stage electrohydraulic control valves. This work is extended in this paper to 2-stage valves of the type shown in Fig. 1. Most of the essential nonlinearities are included in the describing function which is brought to a form similar to that for the single-stage valve.[1] Precalculated charts are utilized which help to make reasonable the amount of labor of computing the describing function for a specific valve. An example in the paper illustrates the use of the describing function, and very good agreement with analog computer studies is demonstrated.

The System Studied

The type of valve shown in Fig. 1 has a second stage which is essentially identical to a single-stage valve except that the control force on the spindle is exerted through a pressure difference, P_1-P_2, at the spindle ends rather than directly through the torque motor. The pressure difference in turn is produced by the first stage, or hydraulic preamplifier, in which the torque motor regulates the relative flow of two leakage orifices, O_1 and O_2 respectively, and thereby creates the pressure difference, P_1-P_2, between the two leakage channels which are fed from the oil accumulator through two fixed orifices, O_3 and O_4 respectively. Special dashpot dampers are usually not incorporated in such valves but the leakage channels and orifices of the first stage exert a dashpot-type damping on the valve spindle.

Nonlinearities Considered

Nonlinearities which were considered in the present study include the square-root flow-pressure relations on all six orifices; the nonlinear reaction forces[2,3] on the valve spindle, *S*, and on the torque motor flapper, *F*; coulomb friction on the valve spindle and flapper; and hysteresis in the torque-motor core. The nonlinearity of the damping is represented by the nonlinearity of the orifice flows. The oil is considered incompressible while metered by the valve, but the oil volumes stored in the actuator and tubing are to be considered compressible although for the small actuator motion involved it is fully justified to treat the compressible oil as a linear spring.

Nomenclature

Any consistent system of units can be used for the following quantities.

A = cross-sectional area of actuator piston
A_c = cross-sectional area of control orifices
A_o = cross-sectional area of fixed orifices
A_s = cross-sectional area of valve spindle
$a, b, \ldots h$ = describing function coefficients; see equations 2 and 3 and Figs. 2–9 of reference 1
$B = 2C_qK_a \cos 69$ deg = reaction (Bernoulli) force coefficient[2,3]
C_q, C_{qc}, and C_{qF} = discharge coefficients for metering orifices, control orifices, and fixed orifices, respectively
D = distance between flapper and either control orifice when flapper is centered
F_c = coulomb friction of spindle
F_{ct} = coulomb friction of flapper
f_m = viscous damping coefficient of second-stage spindle
$G(ju)$ = transfer function of load (see equation 36)
I = torque-motor current
I_h = half-width of torque-motor current hysteresis loop
$i = \dfrac{K_TI}{k_tD}$ = normalized form of I
i_l = linearized torque-motor current
$i_1 = \dfrac{i}{1-4p_ov_t}$, $i_{1l} = \dfrac{i_l}{1-4p_ov_t}$ = modified forms of i and i_l, respectively
$K = K_1K_LK_t$ = open-loop gain of the linearized system
K_a = peripheral width of metering orifices
$K_1G_1(ju)$ = transfer function of linearized valve (equation 33)
$K_LG_L(ju)$ = transfer function of load (equation 35)
$K_tG_t(ju)$ = transfer function of torque-motor electric circuit (equation 41)
$K_v = \sqrt{1/\rho}$ constant of hydraulic fluid
K_T = torque-motor force constant
k_m = spring constant of spindle restoring springs
k_t = spring constant of flapper restoring spring
L_c = circumference of control orifices
l = length of fixed orifices
M_m = effective mass of spindle and restoring springs
m = normalized force coefficient coupling first and second stages
$N(\psi, u)$ = describing function
$P = P_h - P_s$ = accumulator-sump pressure differential
P_a = pressure differential (pressure in right chamber minus pressure in left chamber, see Fig. 1) across actuator piston
P_h = high, or accumulator, pressure to valve
P_s = sump pressure
P_1, P_2 = pressures in left and right chambers, respectively, of first stage
$p = P_a/P$ = normalized form of P_a
$p_h = P_h/P$ = normalized form of P_h
$p_s = P_s/P$ = normalized form of P_s
p_1, $p_2 = P_1/P$, P_2/P = normalized forms of P_1 and P_2, respectively
$\bar{p}_1$ = time average value of p_1 or p_2
p_o = normalized, time-average pressure drop across control orifices (equation 23)
p_{o1} = normalized, time-average pressure drop across fixed orifice (equations 73 and 74)
p_{oo} = modified form of p_{o1} (equation 24)
p_{p1} and p_{p2} = periodic components of p_1 and p_2, respectively
q = normalized flow coefficient of metering orifices (equation 13)
q_c = normalized flow coefficient of control orifices (equation 17)
r_o = radius of fixed orifices
r_1, r_2 = normalized pressure-flow coefficients of the fixed orifices (equations 15 and 16)
$T = C_qK_aK_v$ = flow coefficient of metering orifices
t = time
$u = \dfrac{\omega}{\omega_n}$ = normalized frequency
v = spring nonlinearity ratio (equation 9)
$v_1 = \dfrac{v_t}{1-4p_ov_t}$ = modified form of v_t
v_t = normalized reaction (Bernoulli) force on flapper (equation 18)
X_a = quantity in load monitored for feedback purposes
X_s = valve spindle position from center; positive to right (Fig. 1)
X_{so} = maximum valve spindle position from center
X_0 = idealized actuator-piston displacement for incompressible oil and no oil-ring backlash; positive displacement from center to right (Fig. 1)
X_{00} = maximum actuator-piston displacement
X_t = flapper position; positive from center to left (Fig. 1)

Paper 57-779, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted March 25, 1957; made available for printing July 9, 1957.

J. ZABORSZKY is with Washington University, and consultant to McDonnell Aircraft Corporation, St. Louis, Mo.; and H. J. HARRINGTON is with the McDonnell Aircraft Corporation, St. Louis, Mo.

The authors wish to thank R. C. Cowles for carrying out much of the computations and algebraic operations for the material included in the paper.

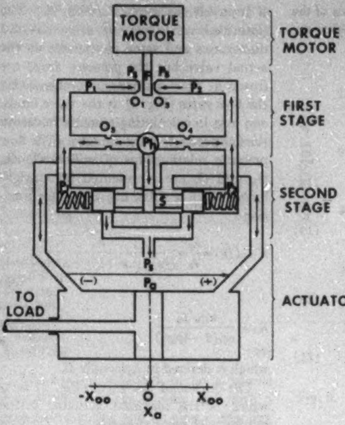

TORQUE MOTOR
P_3
F
P_3
P_1
P_2
O_1
O_2
O_3
O_4
P_h
P
S
P
P_3
(−)
P_a
(+)
TO LOAD
$-X_{00}$
O
X_{00}
X_a
TORQUE MOTOR
FIRST STAGE
SECOND STAGE
ACTUATOR

Fig. 1 (left). Sketch of a 2-stage electro-hydraulic control valve

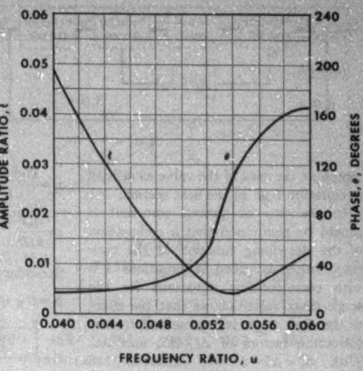


Fig. 3 (right). Frequency response of G(ju) $=\xi/\theta$ for the illustrative example

$x_a=\dfrac{X_a}{X_{oo}}$ = normalized form of X_a

$x_e=\dfrac{X_e}{X_{oo}}$ = normalized form of X_e

$x_0=\dfrac{X_0}{X_{oo}}$ = normalized form of X_0

$x_t=\dfrac{X_t}{D}$ = normalized form of X_t

x_1', x_2' = normalized flow rate of left and right control orifices, respectively

x_3', x_4' = normalized flow rate of left and right fixed orifices, respectively

s = normalized pressure-force coefficient at flapper (equation 19)

$s_1=\dfrac{s+v_t}{1-4p_0v_t}$ = modified form of s (equation 68)

ζ = normalized viscous friction coefficient of spindle (equation 10)

ζ_c = normalized coulomb friction coefficient of spindle (equation 11)

ζ_{ct} = normalized coulomb friction-hysteresis coefficient of flapper and torque motor (equation 12)

$\zeta_{ct1}=\dfrac{\zeta_{ct}}{1-4p_0v_t}$ = modified form of ζ_{ct} (equation 68)

θ = phase angle of $G(ju)$

μ = viscosity of hydraulic fluid

ξ = amplitude of $G(ju)$

ρ = density of hydraulic fluid

$\tau=\omega_n t$ = normalized time

ψ = peak amplitude of sinusoidal oscillation in pressure ratio, p

ω = circular frequency of a sinusoidal oscillation

ω_n = undamped natural frequency of the linearized valve spindle

Summary of the Working Formulas for the Describing Function

The describing function is derived in Appendix I and in dimensionless form is given by:

$$N(\psi, u)=\sqrt{\frac{(1-u^2)^2+4u^2(\zeta+\Delta\zeta)^2}{(C+\Delta C)^2+(S+\Delta S)^2}}\times \angle\, \theta+\tan^{-1}\frac{S+\Delta S}{C+\Delta C}+\tan^{-1}\times \frac{2(\zeta+\Delta\zeta)u}{1-u^2}+90\text{ deg} \tag{1}$$

where

$$C=a(u^2-1)-b2\zeta u+cv\psi-e\frac{2\zeta_c}{q\xi\psi u} \tag{2}$$

$$S=b(u^2-1)+a2\zeta u+fv\psi+h\frac{2\zeta_c}{q\xi\psi u} \tag{3}$$

and

$$\Delta\zeta=\frac{n_1}{2} \tag{4}$$

$$\Delta S=n_1au+n_2\xi^2\psi^2S(C^2+S^2)u^2-\frac{2\zeta_{ct}}{q_c\xi\psi u}\times \frac{n_3C-n_4bu}{\sqrt{C^2+S^2+n_5(bC-aS)u+\frac{1}{4}n_5^2(a^2+b^2)u^2}} \tag{5}$$

$$\Delta C=-n_1bu+n_2\xi^2\psi^2C(C^2+S^2)u^2+\frac{2\zeta_{ct}}{q_c\xi\psi u}\times \frac{n_3S+n_4au}{\sqrt{C^2+S^2+n_5(bC-aS)u+\frac{1}{4}n_5^2(a^2+b^2)u^2}} \tag{6}$$

Equations 5 and 6 assume that terms in ΔS and ΔC which are of order u^3 or higher are negligible which would be true for most practical systems.

The coefficients $a, b, c, e, f,$ and h are the same derived for the one stage valve[1] and can be obtained from the charts of Figs. 2–9 of reference 1. The parameter ψ is the peak value of the actuator pressure differential, p, when it is varying sinusoidally and so ψ establishes the amplitude of possible oscillations. The angle θ is the phase and ξ is the amplitude of the conventional linear transfer function, $G(ju)$, which links the idealized (incompressible oil and no oil-ring backlash) actuator piston motion, x_0, as output to the pressure differential, p, across the actuator as input (see equation 35). The rest of the coefficients in equations 1–6

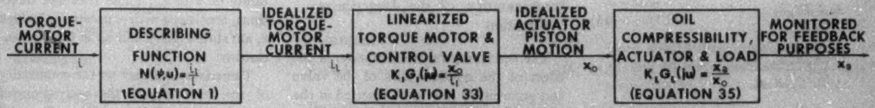

TORQUE-MOTOR CURRENT
i
DESCRIBING FUNCTION
$N(\psi,u)=\frac{i_1}{i}$
(EQUATION 1)
IDEALIZED TORQUE-MOTOR CURRENT
i_1
LINEARIZED TORQUE MOTOR & CONTROL VALVE
$K_1G_1(ju)=\frac{x_0}{i_1}$
(EQUATION 33)
IDEALIZED ACTUATOR PISTON MOTION
x_0
OIL COMPRESSIBILITY, ACTUATOR & LOAD
$K_LG_L(ju)=\frac{x_a}{x_0}$
(EQUATION 35)
MONITORED FOR FEEDBACK PURPOSES
x_a

Fig. 2. Block diagram interpretation of the describing function

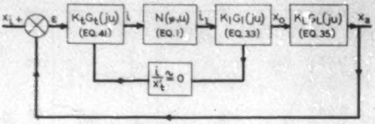


Fig. 4. Block diagram of the system in the illustrative example

represent the data of the valve as defined in equation 7–32 in the next section.

In the way of general comment it should be mentioned that a comparison of the describing function of the two-stage valve as given in equations 1–6 with corresponding equations for the single-stage valve[1] shows that the effect of the first stage is represented by the correction factors of $\Delta\zeta$, ΔS, and ΔC. With $\Delta\zeta=\Delta S=\Delta C=0$, equations 1–6 reduce to those of a single-stage valve with no dashpot damping. The latter is absent because damping is provided by the orifices of the first stage in a double-stage valve.

Nondimensional Parameters

Nondimensional parameters which are directly involved in the computation of the describing function are defined below and can be computed from the actual valve dimensions and system parameters by the following equations:

Primary parameters:

$$\omega_n=\sqrt{\frac{k_m+BP}{M_m}} \tag{7}$$

$$u=\frac{\omega}{\omega_n} \tag{8}$$

$$v=\frac{BP}{k_m+BP} \tag{9}$$

$$\zeta=\frac{f_m}{2\sqrt{M_m(k_m+BP)}} \tag{10}$$

$$\zeta_c=\frac{F_c}{2(k_m+BP)X_{eo}} \tag{11}$$

$$\zeta_{ct}=\frac{F_{ct}+K_TI_h}{2k_tD} \tag{12}$$

$$q=\frac{\omega_nA}{T}\sqrt{\frac{1}{P}\frac{X_{oo}}{X_{eo}}} \tag{13}$$

$$m=\frac{A_vP}{X_{eo}(k_m+BP)} \tag{14}$$

$$r_1=\frac{8\mu A_vlX_{eo}\omega_n}{\pi Pr_o^4} \tag{15}$$

$$r_2=\frac{A_vX_{eo}\omega_n}{\sqrt{2P}C_{qF}K_vA_o} \tag{16}$$

$$q_c=\frac{A_vX_{eo}\omega_n}{\sqrt{2P}C_{qc}K_vL_cD} \tag{17}$$

$$v_t=\frac{8\pi C_{qc}^2PD}{k_t} \tag{18}$$

$$z=\frac{A_cP}{k_tD} \tag{19}$$

Secondary parameters:

$$p_1=G-\sqrt{G^2-F} \tag{20}$$

$$G=\frac{\left[1+\frac{r_2^2}{q_c^2}\right]\left[P_h+\frac{r_2^2}{q_c^2}P_s\right]+\frac{r_1^2}{2q_c^2}P}{P\left[1+\frac{r_2^2}{q_c^2}\right]^2} \tag{21}$$

$$F=\frac{\left[P_h+\frac{r_2^2}{q_c^2}P_s\right]^2+\frac{r_1^2}{q_c^2}P_sP}{P^2\left[1+\frac{r_2^2}{q_c^2}\right]^2} \tag{22}$$

$$p_o=p_1-\frac{P_s}{P} \tag{23}$$

$$p_{oo}=\frac{P_h}{P}+\frac{r_1^2}{4r_2^2}-p_1 \tag{24}$$

$$\nu_1=\frac{3r_1}{2r_2}-\frac{(3p_{oo}-p_o)(p_o+p_{oo})}{p_{oo}\sqrt{p_{oo}}} \tag{25}$$

$$\nu_2=\frac{r_1}{2r_2}-\frac{p_o+p_{oo}}{\sqrt{p_{oo}}} \tag{26}$$

$$\nu_3=-\nu_2+\frac{4r_2p_o\sqrt{p_o}}{q_c}\frac{z+v_t}{1-4p_ov_t} \tag{27}$$

$$n_1=\frac{4mq_cr_2\sqrt{p_op_{oo}}\,(1-4p_ov_t)}{(1-4p_ov_t)(q_c\sqrt{p_o}+r_2\sqrt{p_{oo}})+4r_2p_o\sqrt{p_{oo}}\,(z+v_t)} \tag{28}$$

$$n_2=\frac{q^2\nu_1^2}{2(8mp_o)^2\nu_2\nu_3}\times\left[1+\frac{8p_ov_t}{1-4p_ov_t}\left(\frac{\nu_3}{\nu_1}+\frac{3q_c}{r_2\sqrt{p_o}}\frac{\nu_2^2}{\nu_1^2}\right)\right] \tag{29}$$

$$n_3=-\frac{16mr_2p_o\sqrt{p_o}}{\pi q\nu_3(1-4p_ov_t)} \tag{30}$$

$$n_4=\frac{(8mr_2p_o)^2\sqrt{p_o}}{\pi q\nu_2\nu_3(1-4p_ov_1)} \tag{31}$$

$$n_5=\frac{8mr_2p_o}{\nu_2} \tag{32}$$

Interpretation of the Describing Function

The describing function represents the effect of the nonlinearities of the valve and accordingly it is to be inserted in the block diagram ahead of an equivalent valve which is linearized by depriving it from all nonlinearities (Fig. 2). The linearized valve has the same physical dimensions and same parameters as the actual valve but the pressure drops are frozen at their maximums in calculating the flow rates (Fig. 1) at the valve lands, and also in calculating reaction forces on the flapper and on the spindle. The flow pressure relationships of orifices O_1, O_2, O_3, and O_4 are represented by straight lines. Such a valve has a transfer function:

$$K_1G_1(ju)=\frac{x_o}{i_t}=\frac{K_1}{(ju)[(ju)^2+2(\zeta+\Delta\zeta)(ju)+1]} \tag{33}$$

$$K_1=\frac{n_1\sqrt{p_o}}{qq_c(1-4p_ov_t)} \tag{34}$$

which is derived in Appendix II.

The output of the valve (Fig. 2) is x_0 which is the idealized actuator piston motion with incompressible oil. The load then is represented in the block diagram (Fig. 2) by a transfer function:

$$K_LG_L(ju)=\frac{x_a}{x_o} \tag{35}$$

which has for its output whatever quantity x_a is convenient for this purpose in the block diagram. Usually x_a would be a quantity monitored for feedback purposes such as the actual actuator piston motion.

Another transfer function of the load is needed for computing the describing function (equations 1 through 6),

$$G(ju)=\xi\underline{/\theta}=\frac{x_o}{p} \tag{36}$$

which has the idealized actuator piston motion x_o for its output and the pressure drop across the actuator as its input.

Both of these transfer functions include the compressibility of the oil which, however, is considered a linear spring since the actuator amplitudes are small in the phenomenon considered here. The load for both $G_L(ju)$ and $G(ju)$ is simply defined as the entire dynamic system driven by the actuator including any internal feedbacks but excluding any feedbacks which close through the describing function.

Attention should be paid to the reference directions used for x_o and p in computing $G(ju)=\xi\underline{/\theta}$ to avoid an uncertainly of π in θ. Those reference directions, consistent with equations 1 through 6, are stated (for p and x_0) in the Nomenclature.

There is no limit set on the complexity of the load except the restriction of linearity or linearizability. Transfer functions $G_L(ju)$ and $G(ju)$ are fully conventional linear transfer functions. The use of the describing function is also fully conventional.

Illustrative Example

Data of the system:

A hydraulic valve has the following nondimensional parameters:

$$
\begin{aligned}
q&=428.7\\
\zeta&=0.0189\\
\zeta_e&=0.00574\\
\zeta_{et}&=0\\
v&=0.432\\
q_e&=10.69\\
v_t&=0.2696\\
m&=4.168\\
s&=4.226\\
r_1&=2.598\\
r_2&=23.50\\
p_1&=0.1789
\end{aligned}
\tag{37}
$$

The load is assumed to be a simple inertia with viscous friction and with a restraining spring and an elastic coupling between actuator piston and load inertia. The two transfer functions for this load will be assumed as:

$$K_LG_L(ju)=\frac{x_a}{x_0}=\frac{295.4(ju)^2+1.140(ju)+1}{104.6(ju)^4+91.58(ju)^3+347.7(ju)^2+1.648(ju)+1} \tag{38}$$

$$K_L=1.0 \tag{39}$$

$$G(ju)=\frac{x_a}{p}=\xi\underline{/\theta}=-0.01814\times\frac{104.6(ju)^4+91.58(ju)^3+347.7(ju)^2+1.648(ju)+1}{ju[206.2(ju)^3+181.2(ju)^2+102.3(ju)+1]} \tag{40}$$

The frequency response computed from equation 40 is shown in Fig. 3.

The torque motor is essentially an inductance-resistance circuit with a transfer function of:

$$K_tG_t=\frac{i}{e}=\frac{4.625}{0.5ju+1} \tag{41}$$

$$K_t=4.625$$

The transfer function of the linearized valve can be computed from equations 4, 25, 33, and 34 as

$$\Delta\zeta=7.718 \tag{42}$$

and

$$K_1G_1(ju)=\frac{0.001591}{(ju)[(ju)^2+15.47(ju)+1]} \tag{43}$$

$$K_1=0.001591$$

Block Diagram

The block diagram of the system is shown in Fig. 4. The inner loop represents the feedback caused by the back electromotive force induced in the torque motor by movement of the flapper. This inherently small effect, however, will be neglected here for brevity's sake.

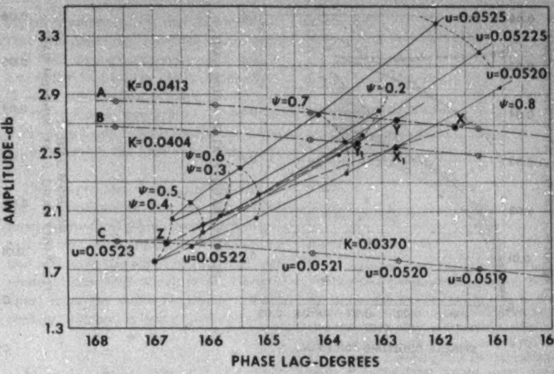


Fig. 5. A sample application of the describing function in the Nichols chart of the illustrative example

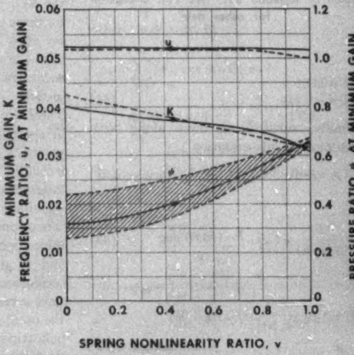


Fig. 6. Effect of the spring nonlinearity ratio, v, on the minimum gain K which is required to sustain oscillations, on the frequency, u, and the pressure amplitude, ψ, of these oscillations. All linear and nonlinear parameters except v have values as given in equation 37. Solid curves represent values obtained by the describing function, dashed curves were obtained by analog computer.[5] There is an uncertainty equal to the shaded area in the analog computer value of ψ

Calculation of the Describing Function

The calculation of one point along the describing function will be illustrated. This will be the point at:

$$u=0.05225 \tag{44}$$

$$\psi=0.4 \tag{45}$$

From Fig. 3 for the selected $u=0.05225$

$$\xi=0.0058 \tag{46}$$

$$\theta=59.83\text{ deg} \tag{47}$$

So from Figs. 2–9 of reference 1 for $\psi=0.4$ and $\theta=59.83$

$$
\begin{aligned}
a&=1.067\\
b&=0.551\\
c&=0.923\\
e&=0.578\\
f&=0.2055\\
h&=1.132
\end{aligned}
\tag{48}
$$

Substituting equation 37 into equations 25 through 32

$$
\begin{aligned}
n_1&=15.44\\
n_2&=6855\\
n_3&=-0.0181
\end{aligned}
\tag{49}
$$

Then substituting equations 37, 44, 45, 48, and 49 into equations 2, 3, 5, and 6, respectively

$$
\begin{aligned}
S&=-0.262\\
C&=-1.034
\end{aligned}
\tag{50}
$$

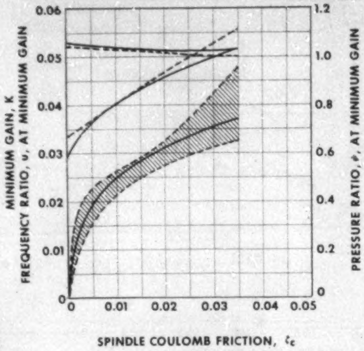


Fig. 7. Effect of the coulomb friction coefficient of the valve spindle, ζ_c, on K, u, and ψ of these oscillations. All linear and nonlinear parameters except ζ_c have values as given in equation 37. See Fig. 6 caption for other data

$$\Delta S = 0.861$$

$$\Delta C = -0.444 \tag{51}$$

Finally, substituting these along with equation 42 into equation 1

$$N(\psi, u) = 0.8056 \underline{/-13.20 \text{ deg}} \tag{52}$$

or

$$-\frac{1}{N(\psi, u)} = 1.241 \underline{/-166.80 \text{ deg}} \tag{53}$$

or

$$20 \log \frac{1}{|N(\psi, u)|} = 20 \log 1.241 = 1.88 \text{ decibels} \tag{54}$$

$$\underline{/-\frac{1}{N(\psi, u)}} = -166.80 \text{ deg} \tag{55}$$

The corresponding point is plotted as point Z in Fig. 5. Other points of the $-1/N(\psi,u)$ curves (the grid of solid and dashed curves in Fig. 5) were obtained in similar fashion. Use has been made of a minor, desk size, digital computer in these calculations. The equations are rather ideally suited for work on such computers.

In accordance with usual describing function[4] and Nichols chart techniques, the $KG_t(ju)\ G_1(ju)\ G_L(ju)$ locus representing the linear part of the open loop is best plotted on a diaphanous sheet on the same scale as the describing function is plotted (Fig. 5). By superimposing this diaphanous sheet on the describing function with ordinate axes lined up, the effect of variations in the gain can be considered by vertically shifting the sheets so that every relative position of the $-1/N(\psi,u)$ loci and the $KG_t(ju)\ G_1(ju)\ G_L(ju)$ locus corresponds to one gain value. In Fig. 5, the $KG_t(ju)\ G_1(ju)\ G_L(ju)$ locus is shown for three gain values; curves A, B, and C for gains of respectively K=0.0413, K=0.0404, and K=0.0370. Curves A and B each show two like-frequency intersections with the $-1/N(\psi, u)$ loci. These are points X at $u=0.05192$ and Y at $u=0.0520$ with gain $K=0.0413$, and points X_1 at $u=0.0520$ and Y_1 at $u=0.05205$ with gain $K=0.0404$. Such intersections indicate the possibility of sustained oscillations.[4] Furthermore, intersections X and X_1 provide feasible sustained oscillations while those at points Y and Y_1 unfeasible ones. The sustained oscillation at points X or X_1 is feasible since a reduction of amplitude, ψ, carries the oscillation below the curves A or B indicating system instability, and so the resulting increasing amplitude returns the system into its oscillation at points X or X_1 respectively. Conversely, an increase of amplitude, ψ, from its value at X shifts the system into the stable region so that the increase is damped out immediately. Consequently points X or X_1 indicate a condition where the system oscillation can be indefinitely sustained. On the other hand, sustained oscillation at points Y or Y_1 is not practically feasible since the crossing of the two curves is reversed at Y relative to the crossing at X. Thus, a repetition of the above discussion shows that, at the points Y or Y_1, either an incipient decrease or an incipient increase of the amplitude becomes cumulative.

Now the behavior of the system for instance with gain $K=0.0404$ can be de-

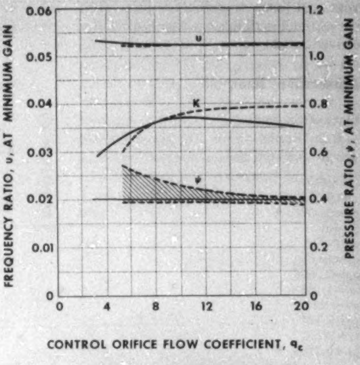


Fig. 8. Effect of the control orifice flow coefficient q_c, on K, u, and ψ of these oscillations. All linear and nonlinear parameters except q_c and $z=1.14/v_t=0.396q_c$ have values as given in equation 37. See Fig. 6 caption for other data

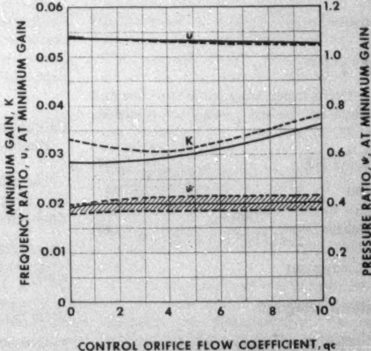


Fig. 9. Effect of the control orifice flow coefficient, q_c, on K, u, and ψ of these oscillations. All linear and nonlinear parameters except q_c and $r_1=0.1105r_2=0.2435q_c$ have values as given in equation 37. See Fig. 6 caption for other data

scribed on the basis of these results of Fig. 5. At this gain the system will not spontaneously start to oscillate; in fact oscillations started by outside disturbances shall die out unless their amplitude exceeds $\psi=0.2$ (point Y_1). If, however, the system is excited to an oscillation of amplitude larger than $\psi=0.2$ this oscillation will either expand or attenuate until its amplitude becomes $\psi=0.73$ (point X_1) where it will be sustained. With changing gain, points X and Y shift along the dashed curves of Fig. 5 (as may be ascertained by plotting intermediate loci of both sets). When the gain is reduced, points X and Y approach each other and finally, at gain $K=0.0370$ (curve C), they merge at point Z. Further reduction of gain results in no more meaningful intersections between the two sets of loci. So $K=0.0370$ is the minimum gain for the given system where sustained oscillation is marginally feasible; below this gain the system is stable in the normal linear sense. This value of gain is shown by a dark dot in Fig. 6 on the curve which shows the minimum gain for sustained oscillation as a function of the nonlinear parameter v. The accompanying amplitude $\psi=0.4$ and the frequency $u=0.05225$ from point Z in Fig. 5 are also shown by dark dots in Fig. 6 on curves of the amplitude and frequency of the minimum-gain oscillation. Other points of the calculated curves in Figs. 7, 8, and 9 were similarly obtained.

The dashed curves in Figs. 6–9 are taken from the analog-computer studies of a companion paper[5] that fully instrumented the nonlinearities of the valve. Comparison of the calculated and measured curves shows very good agreement.

The conditions were such that it was impossible to establish accurate values of the amplitude ψ at the minimum or drop out gain on the analog computer. It was possible, however, to establish the band within which ψ would fall and these bands are shown in Figs. 6–9. The excellent agreement between describing function and analog computer results is evidence of the former's accuracy.

Conclusions

In this paper a describing function has been presented which gives a complete account of the principal nonlinearities of 2-stage electrohydraulic valves, extending previous work for one-stage valves. Precalculated curves reduce the labor for computing the describing function of a specific valve to a very reasonable amount. The necessary computations are well suited for programming on smaller digital computers. It has been shown that this describing function agrees very well with analog computer results and is a powerful tool for the study of the effect of linear and nonlinear valve parameters on the dynamic operation of the valve with given loads. Such studies are helpful in selecting the best valve characteristics.

This paper also makes further use of the generalizations of the describing function method which were introduced earlier[1] for treating multiple nonlinearities and including the effect of the gain and phase of the load which is an element apart from the valve for which the describing function is being developed.

Appendix I

The differential equations of a 2-stage hydraulic valve have been derived in a companion paper;[5] in normalized form they are for the first stage:

$$i=x_t+s(p_1-p_2)+v_t[(1-x_t)^2(p_1-p_s)-(1+x_t)^2(p_2-p_s)]+2\zeta_{ct}\frac{x_t'}{|x_t'|} \qquad (56)$$

$$x_1'=x_3'-x_e' \qquad (57)$$

$$x_2'=x_4'+x_e' \qquad (58)$$

$$g_cx_1'=(1-x_t)\sqrt{p_1-p_s} \qquad (59)$$

$$g_cx_2'=(1+x_t)\sqrt{p_2-p_s} \qquad (60)$$

$$p_h-p_1=r_1x_3'+r_2^2x_3'^2 \qquad (61)$$

$$p_h-p_2=r_1x_4'+r_2^2x_4'^2 \qquad (62)$$

and for the second stage:

$$gx_o'=x_e\sqrt{1+\frac{x_e}{|x_e|}p} \qquad (63)$$

$$m(p_1-p_2)=x_e''+2\zeta x_e'+2\zeta_c\frac{x_e'}{|x_e'|}+x_e+vpx_e \qquad (64)$$

The condition studied here is that of sustained, self-excited or limit-cycle oscillations. If it is assumed that no external signal is introduced into the system then the oscillation is carried out around zero in i, x_e, x_o, and x_t and around a constant average in case of the pressures, p_1 and p_2, in the first stage. It is then profitable to divide the pressure differences into their average and periodic components.

$$p_1=p_1+p_{p1}(t) \qquad (65)$$

$$p_2=p_2+p_{p2}(t) \qquad (66)$$

And since the valve is symmetrically built:

$$p_1=p_2 \qquad (67)$$

Then the first-stage equations 56 through 62 can be written as:

$$i_1=x_t+s_1(p_{p1}-p_{p2})+v_1x_t^2(p_{p1}-p_{p2})-2v_1x_t(p_{p1}+p_{p2})+2\zeta_{ct1}\frac{x_t'}{|x_t'|} \qquad (68)$$

$$x_3'=x_1'+x_e' \qquad (69)$$

$$x_4'=x_2'-x_e' \qquad (70)$$

$$g_cx_1'=(1-x_t)\sqrt{p_0+p_{p1}} \qquad (71)$$

$$g_cx_2'=(1+x_t)\sqrt{p_0+p_{p2}} \qquad (72)$$

$$p_{01}-p_{p1}=r_1x_3'+r_2^2x_3'^2 \qquad (73)$$

$$p_{01}-p_{p2}=r_1x_4'+r_2^2x_4'^2 \qquad (74)$$

where

$$p_0=p_1-p_s=p_2-p_s \qquad (75)$$

$$p_{01}=p_h-p_1=p_h-p_2 \qquad (76)$$

To obtain the describing function, Fourier series representations will be used for i_1, p_{p1}, p_{p2}, x_1', x_2', x_3', x_4', x_e, x_o, and p. Since these functions are unknown beyond the fact that they are periodic, the coefficients of these series will be as yet unknown. It is known however that the two sides of a symmetrically built valve must oscillate exactly 180 deg out of phase with each other. So if

$$p_{p1}=\sum_{n=1}^{\infty}(p_{cn}\cos nu\tau+p_{sn}\sin nu\tau) \qquad (77)$$

then there must be

$$p_{p2}=\sum_{n=1}^{\infty}[p_{cn}\cos n(u\tau+\pi)+p_{sn}\sin n(u\tau+\pi)]=\sum_{n=1}^{\infty}(-1)^n\times(p_{cn}\cos nu\tau+p_{sn}\sin nu\tau) \qquad (78)$$

consequently

$$p_{p1}+p_{p2}=\sum_{n=1}^{\infty}[1+(-1)^n]\,(p_{cn}\cos nu\tau+p_{sn}\sin nu\tau) \qquad (79)$$

$$p_{p1}-p_{p2}=\sum_{n=1}^{\infty}[1-(-1)^n]\times(p_{cn}\cos nu\tau+p_{sn}\sin nu\tau) \qquad (80)$$

and similarly

$$x_1'=\sum_{n=0}^{\infty}(g_n\cos nu\tau+h_n\sin nu\tau) \qquad (81)$$

$$x_2'=\sum_{n=0}^{\infty}(-1)^n(g_n\cos nu\tau+h_n\sin nu\tau) \qquad (82)$$

$$x_3'=\sum_{n=0}^{\infty}(e_n\cos nu\tau+f_n\sin nu\tau) \qquad (83)$$

$$x_4'=\sum_{n=0}^{\infty}(-1)^n(e_n\cos nu\tau+f_n\sin nu\tau) \qquad (84)$$

The definitions of equations 77 through 84 leave only a time shift of π/u to distinguish the counterpair equations in the pairs of respectively 69 and 70, 71 and 72, 73 and 74. So these pairs now represent only one independent equation each and the set of equations 68 through 74 reduces for the first stage to:

$$i_1=x_t+s_1(p_{p1}-p_{p2})+v_1x_t^2(p_{p1}-p_{p2})-$$

$$2v_1x_t(p_{p1}+p_{p2})+2\zeta_{ct_1}\frac{x_t'}{|x_t'|} \qquad (85)$$

$$x_s'=x_1'+x_e' \qquad (86)$$

$$q_cx_1'=(1-x_t)\sqrt{p_o+p_{p1}} \qquad (87)$$

$$p_{o1}-p_{p1}=r_1x_s'+r_2^2x_s'^2 \qquad (88)$$

Eliminating x_s' and x_1' from 86 through 88 with equations 24 and 76

$$x_t=1+\left[\frac{r_1q_c}{2r_2^2}+q_cx_e'\right]\frac{1}{\sqrt{p_o+p_{p1}}}-\frac{q_c}{r_2}\frac{\sqrt{p_{oo}-p_{p1}}}{\sqrt{p_o+p_{p1}}} \qquad (89)$$

Since $p_{p1}<<p_o$ or $p_{p1}<<p_{oo}$, binomial expansion will be used

$$\frac{1}{\sqrt{p_o+p_{p1}}}=\frac{1}{\sqrt{p_o}}\times\left(1-\frac{1}{2p_o}p_{p1}+\frac{3}{8p_o^2}p_{p1}^2-\dots\right) \qquad (90)$$

$$\sqrt{p_{oo}-p_{p1}}=\sqrt{p_{oo}}\times\left(1-\frac{1}{2p_{oo}}p_{p1}-\frac{1}{8p_{oo}^2}p_{p1}^2-\dots\right) \qquad (91)$$

Substituting these into equation 89 and carrying out the multiplications

$$x_t=1+\left(\frac{r_1q_c}{2r_2^2\sqrt{p_o}}+\frac{q_c}{\sqrt{p_o}}x_e'\right)\times\left(1-\frac{1}{2p_o}p_{p1}+\frac{3}{8p_o^2}p_{p1}^2-\dots\right)\times$$
$$-\frac{q_c}{r_2}\sqrt{\frac{p_{oo}}{p_o}}\left(1-\frac{1}{2}\frac{p_o+p_{oo}}{p_op_{oo}}p_{p1}+\frac{3p_{oo}^2+2p_op_{oo}-p_o^2}{8p_o^2p_{oo}^2}p_{p1}^2+\dots\right) \qquad (92)$$

Repeating equation 85

$$i_1=x_t+z_1(p_{p1}-p_{p2})+v_1x_t^2(p_{p1}-p_{p2})-2v_1x_t(p_{p1}+p_{p2})+2\zeta_{ct_1}\frac{x_t'}{|x_t'|} \qquad (93)$$

these two, equations 92 and 93, then represent the first stage of the valve while the second stage is described by equations 63 and 64. The two latter equations have been studied in a previous paper[1] where the assumption of a sinusoidal variation of p was adopted:

$$p=\psi\cos u\tau \qquad (94)$$

On this basis the first, third, and fifth harmonics of

$$x_e=\sum_{n=1,3,5}^{5}(a_n\cos nu\tau+b_n\sin nu\tau) \qquad (95)$$

and the first harmonic of

$$m(p_1-p_2)=m(p_{p1}-p_{p2})=g\xi\psi u[C\cos nu\tau+S\sin nu\tau+\dots] \qquad (96)$$

were determined.[1] The coefficients a_n and b_n, as well as S and C, were found to be functions of the amplitude, ψ, and frequency, u, of the sustained oscillation as well as the phase, θ, and gain, ξ, of the load driven by the actuator. The latter two in turn are again functions of u and ψ.

Consequently in extending this previous study to the first stage of the control valve the a_n, b_n, S and C coefficients can now be assumed known. So comparing equation 96 to equation 80

$$2mp_{c1}=g\xi\psi uC \qquad (97)$$

$$2mp_{s1}=g\xi uS \qquad (98)$$

$$p_{c1}=\frac{g\xi\psi uC}{2m} \qquad (99)$$

$$p_{s1}=\frac{g\xi\psi uS}{2m} \qquad (100)$$

and, retaining only the first two harmonics of p_{p1} and p_{p2}, by equations 77 through 80

$$p_{p1}=\frac{g\xi\psi u}{2m}(C\cos u\tau+C_2\cos 2u\tau+S\sin u\tau+S_2\sin 2u\tau) \qquad (101)$$

$$p_{p1}-p_{p2}=\frac{g\xi\psi u}{m}(C\cos u\tau+S\sin u\tau) \qquad (102)$$

$$p_{p1}+p_{p2}=\frac{g\xi\psi u}{m}(C_2\cos 2u\tau+S_2\sin 2u\tau) \qquad (103)$$

where the second harmonic $C_2=2m\ p_{c2}/g\xi\psi u$ and $S_2=2m\ p_{s2}/g\xi\psi u$ is unknown since it is not directly dependent on the second stage, being absent in $p_{p1}-p_{p2}$, which represents the coupling between the two stages.

Substitution of equations 95 and 101 into equation 92 and application of some well known trigonometric identities results in a form

$$x_t=\sum_{n=0,1,2}[k_n(a_n, b_n, S, C, S_2, C_2)\cos nu\tau+l_n(a_n, b_n, S, C, S_2, C_2)\sin nu\tau] \qquad (104)$$

Now for a symmetrically built valve x_t must not contain second or zero order harmonics since its oscillations will have to be performed symmetrically around zero. So

$$k_0(a_n, b_n, S, C, S_2, C_2)=0 \qquad (105)$$

$$k_2(a_n, b_n, S, C, S_2, C_2)=0 \qquad (106)$$

$$l_2(a_n, b_n, S, C, S_2, C_2)=0 \qquad (107)$$

Equations 106 and 107 can be used to eliminate S_2 and C_2 from k_1 and l_1. Equation 105 could be used for a more accurate establishment of p_1. Then there remains:

$$x_t\cong k_1(a_n, b_n, S, C)\cos u\tau+l_1(a_n, b_n, S, C)\sin u\tau \qquad (108)$$

This now is substituted along with equations 102 and 103 into equation 93, and after further use is made of common trigonometric identities and transformations paralleling those in reference one.

$$i_1=\frac{gq_c}{2\sqrt{p_o}\Delta\zeta}\xi\psi u[(C+\Delta C)\cos u\tau+(S+\Delta S)\sin u\tau+\dots] \qquad (109)$$

is established. The resulting equations for S, C, ΔS, and ΔC are presented in equations 2, 3, 5, and 6 respectively.

A further word is required in connection with handling of the $x_t'/|x_t'|$ term in equation 93. This is a unit amplitude square wave which for a moderately distorted waveform of x_t, as is encountered here, will be essentially in phase with the base harmonic of x_t'. So the base harmonic of the square wave in turn is

$$\frac{x_t'}{|x_t'|}=+\frac{4}{\pi}\times\left[\frac{l_1}{\sqrt{k_1^2+l_1^2}}\cos u\tau-\frac{k_1}{\sqrt{k_1^2+l_1^2}}\sin u\tau\right] \qquad (110)$$

Equation 109 can be written in complex form as

$$i=\frac{\xi\psi u}{K_1}\sqrt{(C+\Delta C)^2+(S+\Delta S)^2}\times\underline{\Big/-\tan^{-1}\frac{S+\Delta S}{C+\Delta C}} \qquad (111)$$

Dividing this into the corresponding equation 121 for the linear valve, equation 1 for the describing function results.

In the preceding derivation the effects of all included nonlinearities on the first harmonic are considered. The effects which are exerted by the higher harmonics on the first through the nonlinearities are included, however the number of these higher harmonics are restricted to the second or the third in most cases.

In conflict with normal describing function practice the output rather than the input is taken as sinusoidal. This approach has been adopted earlier[1] for mathematical expediency and the results here are found to be equally as good as in that previous work.[1]

Appendix II

The linearized form of equations 85 and 88 is defined as respectively

$$i_{1l}=x_t+z_12p_{p1} \qquad (112)$$

$$x_s'=x_1'+x_e' \qquad (113)$$

$$q_cx_1'=(1-x_t)\sqrt{p_o}+\frac{p_{p1}}{2\sqrt{p_o}} \qquad (114)$$

$$x_s'=\frac{1}{2r_2^2}[-r_1+\sqrt{r_1^2+4r_2^2p_{o1}}]-\frac{1}{\sqrt{r_1^2+4r_2^2p_{o1}}}p_{p1} \qquad (115)$$

and for the second stage linearized from equations 63 and 64

$$gx_o'=x_e \qquad (116)$$

$$m2p_{p1}=x_e''+2\zeta x_e'+x_e \qquad (117)$$

Solving equations 113 through 115 and making use of equations 75,76 and 24 and the condition that for the linearized valve p_{p1} is a pure sinusoidal quantity, there result equation 20 and

$$x_t=\frac{q_c}{\sqrt{p_o}}x_e'+\frac{q_c}{\sqrt{p_o}}\times\left[\frac{1}{\sqrt{r_1^2+4r_2^2p_{o1}}}+\frac{1}{2q_c\sqrt{p_o}}\right]p_{p1} \qquad (118)$$

Substituting equation 118 into equation 117 and making use of equations 4 and 116 gives the following:

$$2\Delta\zeta\frac{\sqrt{p_0}}{g\, q_c}\, i_{1l} = x_0''' + 2(\zeta+\Delta\zeta)x_0'' + x_0' \quad (119)$$

Defining

$$K_1 = 2\Delta\zeta\frac{\sqrt{p_0}}{g\, q_c(1-4p_0 v_l)} \quad (120)$$

Equation 119 can be written with $x_0 = \xi\psi\underline{/\theta}$ as

$$i_l = \frac{1}{K_1}\xi\psi u\sqrt{(1-u^2)^2+4(\zeta+\Delta\zeta)^2u^2}$$

$$\underline{/\theta + 90 \text{ deg} - \tan^{-1}\frac{2(\zeta+\Delta\zeta)u}{1-u^2}} \quad (121)$$

and the transfer function of the linearized valve

$$K_1G_1(ju) = \frac{x_0}{i_l} = K_1\frac{1}{ju[(ju)^2+2(\zeta+\Delta\zeta)(ju)+1]} \quad (122)$$

References

1. A Describing Function for Multiple Nonlinearities Present in Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *AIEE Transactions*, vol. 76, pt. I, May 1957, pp. 183–90.

2. Steady-State Axial Forces on Control Valve Pistons, Shih-Ying Lee, J. F. Blackburn. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 74, 1952, pp. 1005–11.

3. Berechnung von Ausfluss und Ueberfallzahlen, R. Von Mises. *Zeitschrift des Vereines Deutscher Ingenieure*, Dusseldorf, Germany, vol. 61, May-June 1917, pp. 447–52.

4. Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements, E. Calvin Johnson. *AIEE Transactions*, pt. II, July 1952, pp. 169–81.

5. Generalized Charts of the Effects of Nonlinearities in 2-Stage Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *Ibid.*, vol. 76, pt. II, 1957 (Jan. 1958 section), pp. 401–08.

6. Generalized Charts of the Effects of Nonlinearities in Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *Ibid.*, vol. 76, pt. I, May 1957, pp. 191–98.

7. A Method for the Selection of Valves and Power Pistons in Hydraulic Servos, F. C. Paddison, W. A. Good. *Report CM-717*, Applied Physics Laboratory, Johns Hopkins University, Baltimore, Md., Jan 26, 1952.

(*Joint discussion with following paper, page 408*)

Generalized Charts of the Effects of Nonlinearities in 2-Stage Electrohydraulic Control Valves

J. ZABORSZKY
MEMBER AIEE

H. J. HARRINGTON
ASSOCIATE MEMBER AIEE

THIS paper, which is an extension of previous work,[1,2] presents a collection of charts derived from full instrumentation on an analog computer of the equations of a closed-loop control system controlled by a 2-stage electrohydraulic control valve of the type shown in Fig. 1. The curves included in this paper show variations in the performance of the control system under the effect of variations of the individual linear and nonlinear parameters of the valve itself and of the other parts of the control system. Usually just one parameter was varied at a time while the other linear and nonlinear components had a constant, nominal value. In some cases, however, as many as three parameters were varied simultaneously. The resulting curves are presented in a general normalized way which permits the use of these results in drawing conclusions on valves of different physical dimensions. As a whole these results are believed to promote a better understanding of the performance of valves of this type. The results listed have also been used to test the accuracy of theoretical investigations.[3]

The System Studied

The study refers to a 2-stage 4-port electrohydraulic valve of the type shown in Fig. 1. The first-stage hydraulic amplifier of this valve regulates by the flapper, F, of the torque motor the relative flow of two orifices O_1 and O_2, and through this device regulates the pressures P_1 and P_2 of two leakage paths, which in turn are being fed by two fixed orifices O_3 and O_4. The P_1-P_2 pressure difference is then used to move the spindle of the second stage of the valve against restoring springs and dynamic forces.

This type of system contains a large number of nonlinearities. The flow relations on all six orifices are quadratic or a combination of linear and quadratic. There are reaction or Bernoulli forces on both the main valve spindle and the flapper of the torque motor. There is coulomb-type friction on the valve spindle and, in some cases, on the flapper; also there is hysteresis in the torque motor. The combination of orifices O_1, O_2, O_3, and O_4 exerts a dashpot-type nonlinear damping on the valve spindle. All these effects, as shown by the equations of Appendixes I and II, were instrumented on the analog computer.

The valve was incorporated in a closed-loop control system as shown in Fig. 2 in block diagram. This system, which also was instrumented on the analog computer, consists of a pure inertia-type load driven by an actuator with elastic coupling between the load and the actuator and a linear restoring spring on the inertia. Viscous and coulomb friction on the actuator and load are assumed, as are additional nonlinear damping terms on the load which are dependent on its acceleration and displacement respectively. The feedback is from the actuator through an electric network. There also is an inner loop resulting from a counter electromotive force induced in the torque motor by the motion of the flapper. The equations for these parts of the system are given in Appendix III. The nominal normalized parameters of the studied system are given in the section of Nominal Parameter Values.

Scope of the Study

The study summarized was concerned with the sustained or limit-cycle oscillations of the system. There are several possible types of oscillations of this kind in the system. Some can be attributed predominantly to a single nonlinearity, for instance coulomb friction on the valve spindle. Others accompany some linear

Paper 57-780, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. Manuscript submitted March 25, 1957; made available for printing July 9, 1957.

J. Zaborszky is with Washington University, and consultant to McDonnell Aircraft Corporation, St. Louis, Mo.; and H. J. Harrington is with McDonnell Aircraft Corporation, St. Louis, Mo.

The authors wish to thank R. C. Cowles for carrying out much of the computations for the material included in this paper.

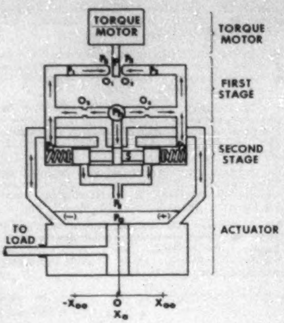


Fig. 1 (left). Sketch of a 2-stage electrohydraulic control valve

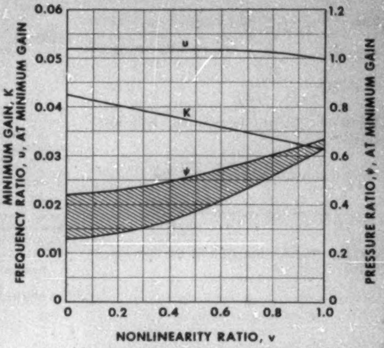


Fig. 3 (right). Effect of the spring nonlinearity ratio, v, on the minimum gain K which is required to sustain oscillations, on the frequency, u, and the pressure amplitude, ψ, of these oscillations. All linear and nonlinear parameters except v have values as given in Nominal Parameter Values section. There is an uncertainty equal to the shaded area in the values of ψ in this curve and in Figs. 4, 15, and 16

resonance peak of the system, such as those around the natural frequency of the load inertia or of the valve spindle inertia. The latter usually show the influence of a large number of nonlinearities simultaneously.

The present study is mainly concerned with the oscillations around the resonant frequency of the load. This appears to be the most significant of the oscillations; not so much because it is highly sensitive to most nonlinear effects simultaneously, but because this tends to be the condition setting the stability margin for a system of this kind. Accordingly the study was directed mainly at finding the minimum or drop out gain, that is the smallest gain which makes it possible to maintain an oscillation, near the resonant frequency of the load. Variations of this gain with various linear and nonlinear parameters were determined. The frequency and amplitude of these sustained oscillations at the drop out gain was also established, these being functions of the gain in nonlinear systems.

The following sections discuss the effects of the various linear and nonlinear parameters in connection with the charts. The information is believed to be conveyed most effectively by the charts which will be commented on briefly.

Effects of the Various Linear and Nonlinear Parameters

EFFECT OF REACTION (BERNOULLI) FORCES ON THE VALVE SPINDLE

The parameter v expresses what portion of the total spring constant on the valve spindle is derived from the reaction or Bernoulli forces[4,5] caused by the flow through the valve with zero load on the actuator piston. This reaction force is proportional to the valve opening; in this it acts as a linear spring. However, it is also proportional to the pressure drop across the valve, and this latter in effect represents a nonlinear feedback to the valve spindle. Fig. 3 shows the effects of v on the valve performance. Increasing v may be seen to exert a strong destabilizing effect, a strong reduction of frequency, u, and also a strong increase of amplitude ψ. Nevertheless the effect of v appears to be much less on the 2-stage valve than on the single-stage one.[1,2] This may be a result of the damping action of the first-stage orifices.

EFFECT OF COULOMB FRICTION AT VARIOUS PARTS OF THE SYSTEM

Coulomb friction at any part of the system exerts a marked effect on the performance. Figs. 4–9 show the effect of coulomb friction at the valve spindle, the actuator, and the load. It should be noted that in all of these coulomb friction exhibits a strong stabilizing effect (an increasing minimum gain for sustained oscillations) with very little effect on the frequency, u, and amplitude, ψ, of these oscillations, except when the coulomb friction is on the valve spindle. In the latter case the amplitude increases substantially with increasing amounts of coulomb friction and drops to zero at no coulomb friction. This does not mean that nonlinear oscillations are not possible without coulomb friction; however the nature of such oscillations changes. With coulomb friction, at the nominal parameters the oscillation is of the hard type, meaning that there is a drop-out gain which maintains an oscillation with a finite amplitude, but upon a slight reduction of this drop-out gain the oscillations disappear altogether. Without coulomb friction, at the nominal parameters the oscillation is of the soft type: the oscillation would start at a certain drop-out gain value with negligible amplitude and the amplitude would gradually increase with increasing gain. Theoretical studies based on describing functions[1,3] indicate that a drop-out gain of the discontinuous or hard type will be present without

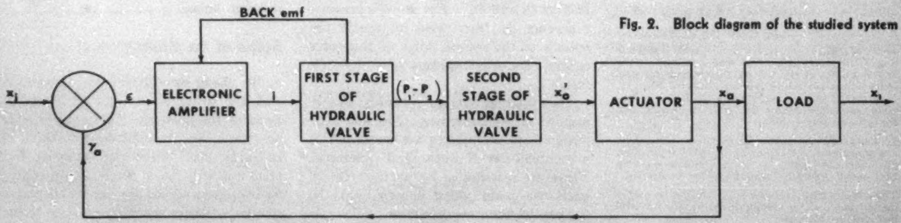


Fig. 2. Block diagram of the studied system

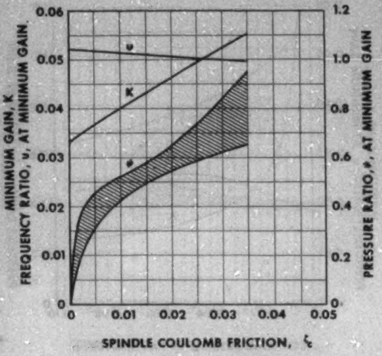
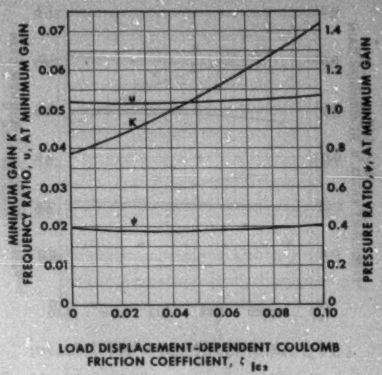


Fig. 4. Effect of the coulomb-friction coefficient of the valve spindle, ζ_c, on K, u, and ψ as in Fig. 3. All parameters except ζ_c are given in values section

Fig. 7 (right). Effect of the displacement dependent coulomb friction coefficient, ζ_{lc2}, on K, u, and ψ as in Fig. 3. All parameters except ζ_{lc2} and $\zeta_{lc1}=\zeta_{lc3}=0$ are given in values section



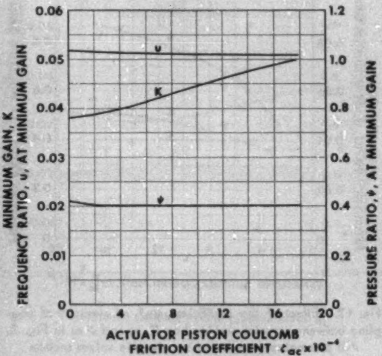


Fig. 5 (left). Effect of the coulomb friction on the actuator piston, ζ_{ac}, on K, u, and ψ as in Fig. 3. All parameters except ζ_{ac} are given in values section

Fig. 8 (right). Effect of the acceleration dependent coulomb friction coefficient of the load, ζ_{lc3}, on K, u, and ψ as in Fig. 3. All parameters except ζ_{lc3} and $\zeta_{lc1}=\zeta_{lc2}=0$ are given in values section

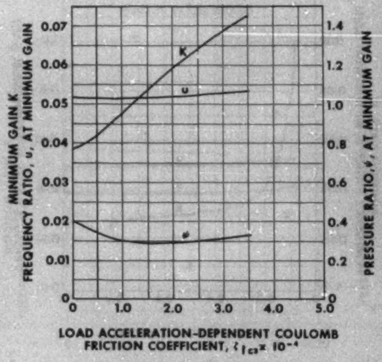
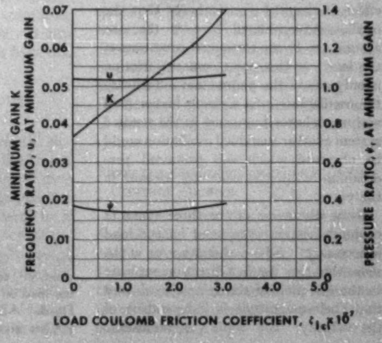


Fig. 9 (below). Effect of the simultaneous variations of the three types of load coulomb friction coefficient ζ_{lc1}, ζ_{lc2} and ζ_{lc3} on K, u, and ψ as in Fig. 3. All parameters except ζ_{lc1} and $\zeta_{lc2} = 403\ \zeta_{lc3} = 1.74 \times 10^3\ \zeta_{lc1}$ are given in values section



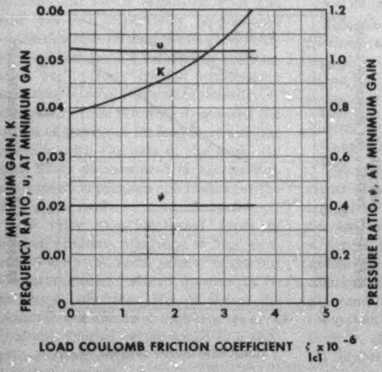


Fig. 6 (left). Effect of the coulomb friction coefficient of the load, ζ_{lc1}, on K, u, and ψ as in Fig. 3. All parameters except ζ_{lc1} and $\zeta_{lc2}=\zeta_{lc3}=0$ are given in values section

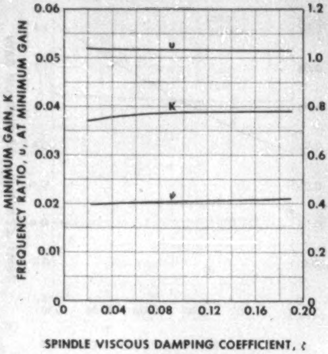
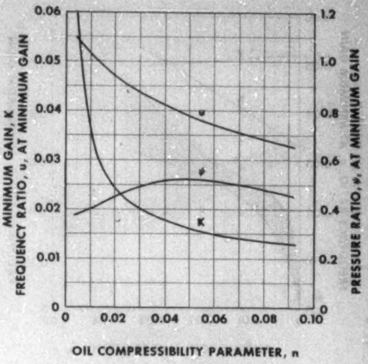


Fig. 10 (left). Effect of the viscous friction on the valve spindle, ζ, on K, u, and ψ as in Fig. 3. All parameters except ζ are given in values section

Fig. 12 (right). Effect of the oil compressibility, n, on K, u, and ψ as in Fig. 3. All parameters except n are given in values section



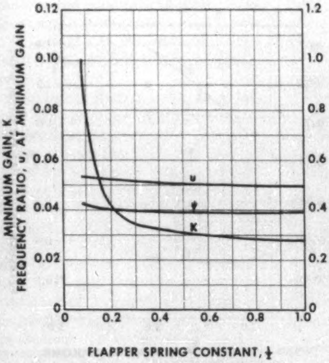


Fig. 11 (left). Effect of the flapper spring constant, 1/z, on K, u, and ψ as in Fig. 3. All parameters except 1/z and $v_l = 0.0638z$ are given in values section

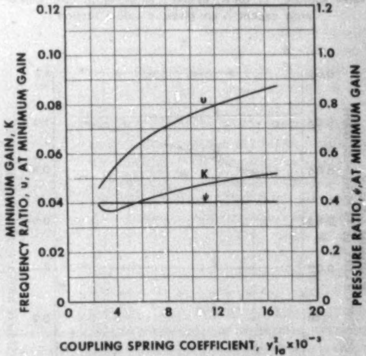


Fig. 13. Effect of the coefficient, y_{1a}^2, of elasticity of coupling between actuator and load on K, u, and ψ as in Fig. 3. All parameters except y_{1a}^2 are given in values section

coulomb friction also if the parameter v exceeds substantially the value of the v in the Nominal Parameter Values section.

If it appears to be a paradox that coulomb friction should have a stabilizing effect, it should be remembered that the oscillation considered here is the one around the linear resonance condition of the load; at this point coulomb friction is only one of the parameters influencing the oscillations. At a much higher gain and much lower frequency in the studied system another sustained oscillation could be observed which was produced predominantly by the coulomb friction in the valve.

Some discussion of the three types of coulomb friction considered on the load is necessary. All are coulomb type in the sense that they act in a direction opposite to the direction of the velocity, x_l', and that their magnitude is independent of the velocity. But while $\zeta_l c_l$ represents

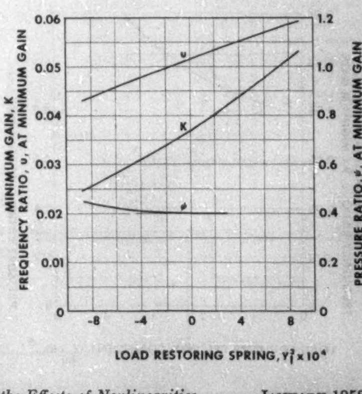


Fig. 14. Effect of the coefficient, y_l^2, of restoring spring of the load on K, u, and ψ as in Fig. 3. All parameters except y_l^2, are given in values section

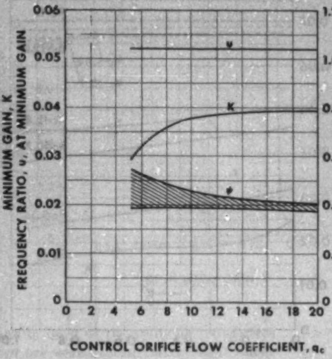


Fig. 15 (left). Effect of the control orifice flow coefficient, q_c, on K, u, and ψ as in Fig. 3. All parameters except q_c and $z = 1.14/v_t = 0.396 \times q_c$ are given in values section

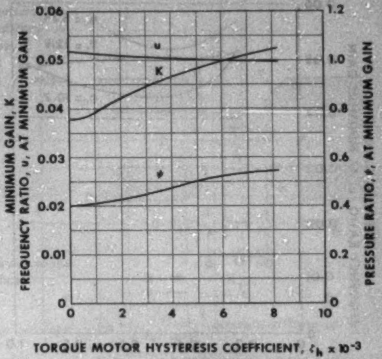


Fig. 17. Effect of the torque motor hysteresis coefficient, ζ_h on K, u, and ψ as in Fig. 3. All parameters except ζ_h are given in values section

Fig. 16 (below). Effect of the control orifice flow coefficient q_c, on K, u, and ψ as in Fig. 3. All parameters except q_c and $r_1 = 0.1105\ r_2 = 0.2435 q_c$ are given in values section

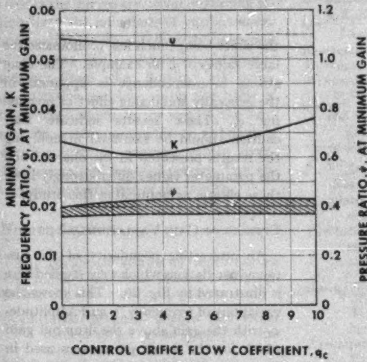
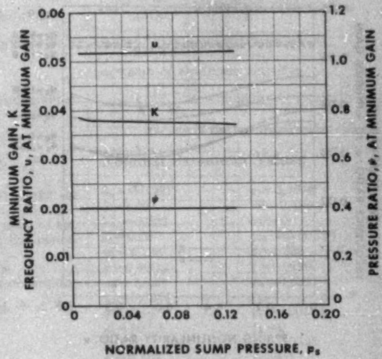


Fig. 18 (right). Effect of the normalized sump pressure, p_s on K, u, ψ as in Fig. 3. All parameters except p_s are given in values section



a conventionally interpreted coulomb friction, the forces associated with ζ_{1c_2} and ζ_{1c_3} are respectively proportional to the absolute value of the displacement, $|x_1|$, and the absolute value of the acceleration, $|x_1''|$. Such friction components may arise on some loads when respectively the center of the restoring torque and the center of gravity of the load inertia do not coincide with the axis of the load. In such cases displacement or acceleration shall result in pressure and friction at the hub of the load.

All three afore-mentioned types of coulomb friction, according to Figs. 6, 7, and 8 have the same general effect but the acceleration induced appears to be the most effective. Fig. 9 shows the simultaneous effects of the three types of coulomb friction terms mixed in a definite arbitrary ratio.

EFFECT OF VISCOUS FRICTION AT THE VALVE SPINDLE

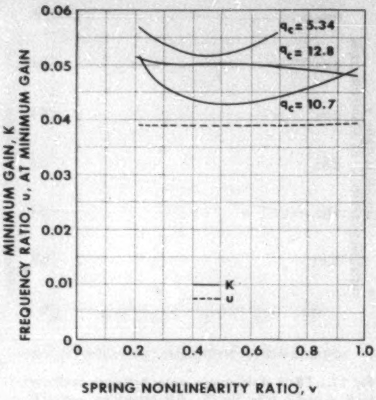
According to Fig. 10 the effect of viscous friction of the valve spindle is a stabilizing one; a predictable result. Fig. 10 also shows, however, the slight effect from this source.

EFFECT OF THE VARIOUS ELASTIC ELEMENTS OF THE SYSTEM

The effects of the various elastic elements within the system are illustrated by Figs. 11–14. From these figures it may be observed that harder springs (smaller n, larger y^2_{1a} or y^2_1) always lead to a marked increase in the frequency of oscillations and, except for small values of y^2_{1a}, tend to have a stabilizing effect on the system. The amplitude of the oscillations is affected but slightly. Finally, the flapper spring constant $1/z$, is effective mainly if it is quite soft; then softer spring means more stability. The effec, of the flapper spring constant on fret quency or on amplitude is always small-however. The physical significance of decreasing the flapper spring constant is to emphasize the importance of the pressure and Bernoulli forces acting on the flapper; apparently these forces act as a stabilizing dynamic feedback for the system.

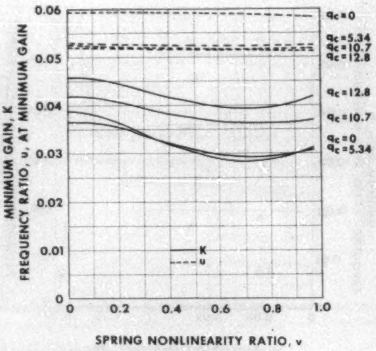
EFFECT OF THE FLOW COEFFICIENT OF THE CONTROL ORIFICE

The effect of the flow coefficient of the control orifice was studied under two conditions as recorded in Figs. 15 and 16. The former chart signifies variations in the dimensional parameter, D, while the latter allows an over-all change in first-stage flow rates.



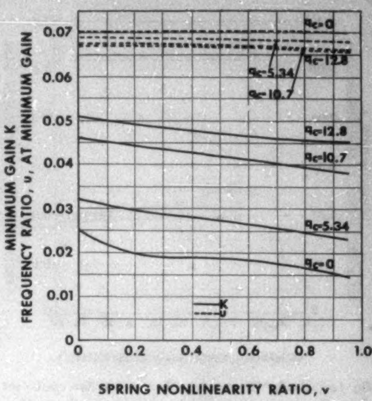
(A)

Fig. 19(A). Effect of the spring nonlinearity ratio, v, and of the flow coefficient, q_c, on K, u, and ψ as in Fig. 3. All parameters except v, q_c, $r_1=0.1105$ $r_2=0.2435q_c$, and $y_{1a}^2 = 0.00169$ are given in the values section



(B)

Fig. 19 (B). Effect of the spring nonlinearity ratio, v, and flow coefficient, q_c, on K, u, and ψ as in Fig. 3. All parameters except v, q_c, $r_1=0.1105 \times r_2=0.2435$ q_c, and $y_{1a}^2=0.00338$ are given in the values section



(C)

Fig. 19(C). Effect of the spring nonlinearity ratio, v, and flow coefficient, q_c, on K, u, and ψ as in Fig. 3. All parameters except v, q_c, $r_1=0.1105$ $r_2=0.2435q_c$, and $y_{1a}^2 = 0.00677$ are given in the values section

Effect of the Hysteresis in the Torque Motor

This effect is shown in Fig. 17 and is found to be somewhat stabilizing at the mode of oscillation near the resonant frequency of the load. In this it is similar to the effects of coulomb friction in various parts of the system as should be expected.

Effect of Normalized Sump Pressure

The effect of sump pressure is quite negligible according to Fig. 18, at least within the studied range over which this pressure has been varied.

Simultaneous Variation of Three Parameters

The three parameters selected for variation in obtaining Figs. 19(A), 19(B), and 19(C) are

1. The spring nonlinearity ratio, v, of the valve which has a rather dominating effect on the performance.
2. The stiffness, y_{la}^2, of the linkages between actuator piston and load inertia which determines primarily the frequency of the oscillations.
3. The flow parameter, q_c, of the control orifice.

It is interesting that with the two stage valve the effect of v, the spring nonlinearity ratio, is highly dependent on the parameters y_{la}^2 and q_c. In fact even the tendency of variation with v can be reversed; with some q_c and y_{la}^2 combinations it is stabilizing, with others it is destabilizing. Increasing y_{la}^2 is generally destabilizing and results in increasing the frequency, u. However the destabilizing effect of y_{la}^2 is also strongly dependent on the other two parameters, the effect being much less pronounced for large values of q_c for example. The same strong interdependence is apparent in the generally stabilizing effect of increasing q_c. These results indicate that caution should be exercised in applying the results presented in the charts when the parameter values differ strongly from those chosen as nominal in this study.

Effects of Gain Variations

An interesting peculiarity of the sustained oscillations which were studied here is illustrated by Fig. 20. This shows the variations of frequency, u and amplitude-ψ, with the gain above the drop-out gain which for the set of parameters used in Fig. 20 is $K = 0.04$. Above this gain value the oscillations diverge rather rapidly in amplitude. However stable oscillations are again resumed at about $K=0.18$ and these latter oscillations converge with increasing gain. The frequency is near the natural frequency of the load for both oscillations. A loop in the Nichols chart of the system together with describing function considerations[3] indicates the presence of these two stable oscillating regions.

Conclusions

The results of the analog computer study of 2-stage electrohydraulic control valves is presented. It would be a rather impossible task to vary all the numerous linear and nonlinear parameters simultaneously over their practical ranges. Consequently only a maximum of three parameters has been varied in this fashion,

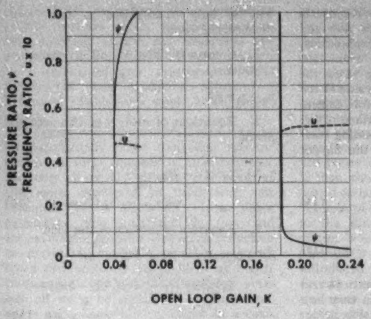


Fig. 20 (left). Loci of frequency, u, and pressure amplitude, ψ, of sustained oscillations versus gain, K. All parameters except y_{la}^2 = 0.00254 are given in values section

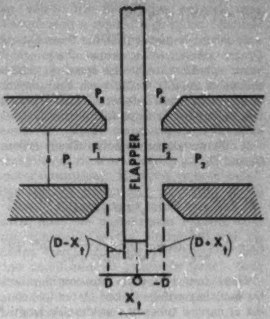


Fig. 21 (right). Sketch of the control orifice

and in most of the study all parameters but one were kept at the constant values given in the section of Nominal Parameter Values. Even so it is believed that the results of the study as summarized in the curves are useful in gaining insight into the influences of various linear and nonlinear parameters on the operation of systems with two stage hydraulic valves. Even the drawing of semiquantitative conclusions on the basis of the curves seems fairly safe if the normalized system parameters do not deviate very far from the nominal values given. The normalized form of presentation enhances the usefulness of the results.

Nomenclature*

Any consistent system of units can be used for the quantities defined in the following.

D = distance between flapper and either control orifice when flapper is centered; see Fig. 21

$E = X_t - K_f X_0$ = servo error

F_{ac} = coulomb-friction force of actuator piston

f_a = viscous damping coefficient of actuator piston

F_{lc1} = coulomb-friction force of load

F_{lc2} = coefficient of amplitude-dependent coulomb friction of load

F_{lc3} = coefficient of acceleration-dependent coulomb friction of load

f_l = viscous damping coefficient of load

$h_a = \dfrac{AP}{M_a X_{oo} \omega_n^2}$ = normalized form of A

$[I]_h$ = torque motor current distorted by hysteresis

$[i]_h = \dfrac{K_T}{Dk_t}[I]_h$ = normalized form of $[I]_h$

$k = \dfrac{X_{oo}}{X_{lo}}$ = proportionality factor

K_{at} = coefficient of back electromotive force (emf) induced by motion of flapper in magnetic field

$k_{at} = \dfrac{K_{at} K_T}{k_t L}$ = normalized form of K_{at}

K_f = electrical gain of feedback transducer; see E

K_{la} = spring constant of linkage between actuator piston and load

K_T = force constant of torque motor

k_t = equivalent spring constant of flapper restoring spring (including magnetic forces)

L = inductance of torque motor coil

$l_a = \dfrac{\mu_a K_f K_T X_{oo}}{D k_t L \omega_n}$ = normalized form of μ_a

K_l = restoring spring constant of load

M_a = mass of actuator piston

M_l = mass of load

$n = \dfrac{V_o P}{4\beta A X_{oo}}$ = normalized oil compressibility parameter

R = resistance of torque motor coil

$r = \dfrac{R + r_p}{L\omega_n}$ = normalized form of $R + r_p$

r_p = plate resistance of output stage of electronic amplifier

V_o = total volume of oil on both sides of the actuator piston between the piston and the valve

$w = \sqrt{\dfrac{K_{la}}{M_a \omega_n^2}}$ = normalized form of the undamped natural frequency of the actuator piston

X_t = input emf to the servo

$x_t = \dfrac{X_t}{K_f X_{oo}}$ = normalized form of X_t

X_l = displacement of the load

X_{lo} = maximum displacement of load from center

$x_l = \dfrac{X_l}{X_{lo}}$ = normalized form of X_l

$y_l = \sqrt{\dfrac{K_l}{M_l \omega_n^2}}$ = normalized form of the undamped natural frequency of the load referred to K_l when K_{la} is assumed zero

$y_{la} = \sqrt{\dfrac{K_{la}}{M_l \omega_n^2}}$ = normalized form of the undamped natural frequency of the load referred to K_{la} when K_l is assumed zero

β = bulk modulus of the oil

δ = diameter of control orifice; see Fig. 21

$\epsilon = \dfrac{E}{K_f X_{oo}}$ = normalized form of E

$\zeta_a = \dfrac{f_a}{2\sqrt{M_a K_{la}}}$ = normalized form of f_a

$\zeta_{ac} = \dfrac{F_{ac}}{2 M_a X_{oo} \omega_n^2}$ = normalized form of F_{ac}

$\zeta_h = \dfrac{K_T I_h}{2 k_t D}$ = normalized hysteresis coefficient

$\zeta_{lc1} = \dfrac{F_{lc1}}{2 M_l X_{lo} \omega_n^2}$ = normalized form of F_{lc1}

$\zeta_{lc2} = \dfrac{F_{lc2}}{2 M_l \omega_n^2}$ = normalized form of F_{lc2}

$\zeta_{lc3} = \dfrac{F_{lc3}}{2 M_l}$ = normalized form of F_{lc3}

μ_a = gain of electronic amplifier

$\omega_n = \sqrt{\dfrac{k_m + BP}{M_m}}$ = undamped natural frequency of the linearized valve spindle

Nominal Parameter Values

$h_a = 0.0260$
$k = 1.00$
$k_{at} = 0.322$
$l_a = 9.26$
$m = 4.168$
$n = 0.00920$
$p_h = 1.023$
$p_s = 0.0234$
$q = 428.7$
$q_c = 10.69$
$r = 2.0$
$r_1 = 2.598$
$r_2 = 23.50$
$v = 0.432$
$v_t = 0.2696$
$w^2 = 0.489$
$y_l^2 = 0$
$y_{la}^2 = 0.00338$
$z = 4.226$
$\zeta = 0.0190$
$\zeta_a = 0.626$
$\zeta_{ac} = 0$
$\zeta_c = 0.00574$
$\zeta_h = 0$
$\zeta_l = 0.0333$
$\zeta_{lc1} = 0$
$\zeta_{lc2} = 0$
$\zeta_{lc3} = 0$

Appendix I. Derivation of the Equations of the First Stage of the Valve

The various equations shall be written first in dimensional form and then normalized.

1. Flow relations on the two control orifices O_1 and O_2

$$Q_1 = \sqrt{2}\, C_{qc} K_v L_c (D - X_t) \sqrt{P_1 - P_s} \qquad (1)$$

$$Q_2 = \sqrt{2}\, C_{qc} K_v L_c (D + X_t) \sqrt{P_2 - P_s} \qquad (2)$$

These are simply the general flow relations for sharp-edged orifices (assuming $D << \delta$) with $P_1 - P_s$ and $P_2 - P_s$ being the respective pressure drops across the orifices and $L_c(D - x_t)$ and $L_c(D + x_t)$ the respective orifice cross sections (Fig. 21). Normalized, equations 1 and 2 become

$$q_c x_1' = (1 - x_t) \sqrt{p_1 - p_s} \qquad (3)$$

---

* Additional notation may be found in a companion paper.[1]

$$q_c x_3' = (1+x_t)\sqrt{p_3 - p_s} \tag{4}$$

It should be noted that the flows Q_1 and Q_3 are normalized in terms of equivalent valve spindle speed; the speed at which the valve spindle would push the amount of flow in question before it. This principle will be applied in what follows to all flows in the first stage.

2. Flow relations on the fixed orifices O_3 and O_4.

$$P_h - P_1 = \frac{8\mu l}{\pi r_o^4} Q_3 + \frac{1}{(\sqrt{2}\, C_{qF} K_v A_o)^2} Q_3^2 \tag{5}$$

$$P_h - P_2 = \frac{8\mu l}{\pi r_o^4} Q_4 + \frac{1}{(\sqrt{2}\, C_{qF} K_v A_o)^2} Q_4^2 \tag{6}$$

These equations take into consideration the fact that orifices O_3 and O_4 tend to consist of narrow tubes of considerable length. Accordingly the first terms on the right give the pressure drop for the laminar flow in these narrow tubes, while the second terms are the pressure drops caused by the discharge at the tube ends. Normalized, equations 7 and 8 become

$$p_h - p_1 = r_1 x_3' + r_2^2 x_3'^2 \tag{7}$$

$$p_h - p_2 = r_1 x_4' + r_2^2 x_4'^2 \tag{8}$$

3. Continuity relations for the two leakage channels:

$$Q_3 - Q_1 = A_s \dot{X}_e \tag{9}$$

$$Q_4 - Q_2 = -A_s \dot{X}_e \tag{10}$$

which simply expresses the fact that the difference of the oil volume flowing through either pair of fixed and control orifices in series is used to replace oil behind the spindle when the latter is moving. The oil is considered incompressible in these passages which is justified by the small volume and rather uniform pressure. In normalized form, equations 9 and 10 become

$$x_3' = x_1' + x_e' \tag{11}$$

$$x_4' = x_2' - x_e' \tag{12}$$

4. Dynamics of the torque motor flapper:

$$K_T[I]_h = k_t X_t + A_c(P_1 - P_2) + (F_1 - F_2) \tag{13}$$

On the left-hand side stands the electric torque delivered by the torque motor to the flapper, $[I]_h$ being the torque motor current as modified by the hysteresis. On the right the first term is the mechanical spring action on the flapper, which may be its own elasticity, the second term is the force on the flapper which results from the static pressure difference in the two leakage channels. The last term finally represents the reaction forces from the fluid flow and this term merits a few words of explanation. These reaction forces derive from the change of momentum occurring when the fluid that is coming in along the axis of the channel leading to the control orifices discharges radially along the surface of the flapper (Fig. 21). The corresponding reaction force on the two sides of the flapper will be:

$$F_1 = \rho Q_1 V_1 = \rho Q_1^2 \frac{1}{A_c} \tag{14}$$

$$F_2 = \rho Q_2 V_2 = \rho Q_2^2 \frac{1}{A_c} \tag{15}$$

Substituting equation 1 and 2 into 14 and 15, respectively, and the latter in turn into equation 13 there results after normalizing:

$$[i]_h = x_t + z(p_1 - p_2) + v_t[(1-x_t)^2(p_1 - p_s) - (1+x_t)^2(p_1 - p_2)] \tag{16}$$

Appendix II. Equations of the Second Stage of the Valve

These equations have been derived[1,2] with a slight difference and they are in normalized form:

$$m(p_1 - p_2) = x_e'' + 2\zeta x_e' + 2\zeta_e \frac{x_e'}{|x_e'|} + x_e + vp|x_e| \tag{17}$$

and

$$q x_o' \sqrt{1 + p\frac{x_e}{|x_e|}} \tag{18}$$

One difference with reference 1 or 2 is the left hand side of equation 17 which for the one stage valve[1,2] is the force exerted by the torque motor on the valve spindle, and for the 2-stage valve of course is the force exerted by the static pressures at the two ends of the spindle. The other difference is the absence in equation 17 of an explicit dashpot-type damping term, since dashpot damping also is now implicit in $(p_1 - p_2)$.

Appendix III. Differential Equations Exclusive of Control Valve

With the exception of the acceleration- and position-dependent coulomb friction terms of the load, the equations given here are identical to those derived in reference 2. Consequently, the equations will be only summarized in their nondimensional form.

1. Compressibility of oil in actuator cylinder

$$x_a - x_o = np \tag{19}$$

2. Equations of motion of the actuator piston

$$x_a'' + 2\zeta_a w x_a' + 2\zeta_{ac}\frac{x_a'}{|x_a'|} + w^2(x_a - x_l/k) + h_a p = 0 \tag{20}$$

3. Equation of motion of the load

$$x_l'' + 2\zeta_l y_{la} x_l' + [2\zeta_{lc1} + 2\zeta_{lc2}|x_l| + 2\zeta_{lc3}|x_l''|]\frac{x_l'}{|x_l'|} + y_{la}^2(x_l - kx_a) + y_l^2 x_l = 0 \tag{21}$$

Here it is assumed that coulomb friction varies directly with bearing normal force loading and that the latter varies linearly with both the absolute value of load position and of load acceleration.

4. Equation of error signal

$$\epsilon = x_t - x_a \tag{22}$$

5. Differential equation of amplifier and torque motor

$$_a\epsilon = i' + ri + k_a x_t' \tag{23}$$

This equation differs from its counterpart in reference 2 only in that the back emf term is now derived from flapper rate, x_t', rather than spindle rate, x_e'.

References

1. A Describing Function for Multiple Nonlinearities Present in Electrohydraulic Control Valves, **J. Zaborszky, H. J. Harrington.** *AIEE Transactions*, vol. 76, pt. I, May 1957, pp. 183–90.

2. Generalized Charts of the Effects of Nonlinearities in Electrohydraulic Control Valves, **J. Zaborszky, H. J. Harrington.** *Ibid.*, pp. 191–98.

3. A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves, **J. Zaborszky, H. J. Harrington.** *Ibid.*, pt. II, 1957 (Jan. 1958 section), pp. 394–401.

4. Steady-State Axial Forces on Control Valve Pistons, **Shih-Ying Lee, J. F. Blackburn.** *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 74, 1952, pp. 1005–11.

5. Berechnung von Ausfluss und Ueberfallzahlen, **R. Von Mises.** *Zeitschrift des Vereines Deutscher Ingenieure*, Düsseldorf, Germany, vol. 61, May-June 1917, pp. 447–52.

◆

Joint Discussion of Papers 57-779 and 57-780

L. R. Axelrod and W. L. Kinney (Cook Research Laboratories, Morton Grove, Ill.): The authors of this paper and its companion paper are to be complimented for their efforts in applying describing function techniques to the multiple nonlinearities present in an electrohydraulic servo valve. Our organization has also been studying this type of valve, but the analytical aspects of the study have been principally centered around use of small perturbation techniques to establish suitable transfer functions for these valves. This work has been done during the past 4 years under various Air Force contracts and has led to a successful method of predicting dynamic response from system static characteristics. Our present program consists in part of simulating an electrohydraulic servo valve as a function of its internal operating characteristics on an analogue computer. Hence, this paper appeared at a very opportune time.

The authors' use of dimensionless (normalized) quantities was, from our standpoint, an unfortunate choice since the effect on valve operation as caused by a change in one physical parameter became somewhat obscured. Also, since no numerical values for the physical parameters were given it was impossible to compare the authors' normalized servo system to typical systems which we have used to experimentally verify our analytical approach to this problem. If the author's have attempted any experimental verification of their analytical solution, a discussion of this work would be enlightening.

We agree with the author's statement that electrohydraulic valve controlled systems contain a large number of nonlinearities. However, it is our opinion that some of the more significant of these were not included. Items such as torque-motor saturation, flow saturation in first and second stages, pressure saturation, reduction in maximum available pressure at small spool displacements (a problem at higher frequencies), and amount of under or overlap all may be significant in an actual valve and should be considered in a complete analysis.

J. Zaborszky and H. J. Harrington: The comments of Messrs. Axelrod and Kinney are appreciated. It would seem, however, that the use of a normalized treatment should facilitate rather than hinder the comparison of the results of this paper and its companion[1] to their physical systems. For this sort of comparison, of course, it is necessary to compute the normalized parameters from the physical parameters of their system. This can be accomplished with the set of equations 7 through 32 of the first paper[1] and the notations of both papers. If then the normalized parameters so obtained do not match those of the illustrative example of the paper, results corresponding to the illustrative example, but for the system in question, can be obtained through comparatively little effort with the methods introduced in the papers. In other words, the use of the nondimensional or normalized treatment makes the results universally comparable to any physical system. On the other hand, if the analysis of the paper should have been carried out for a specific physical system, then comparison to physical systems of differing parameters would be impossible. The physical system for which the illustrative example was worked used a typical high-speed 5-gallon-per-minute 2,000-pound-per-square-inch valve, and experimental studies gave equally as good agreement with the theoretical work as that indicated in earlier papers[2,3] for the single-stage valve.

It appears to the authors that all of the nonlinear effects mentioned by the discussers are included in the papers with the exception of over or underlap and of torque motor saturation. However, over or underlap are not universally significant in all valves. Torque motor saturation does not enter into the phenomenon studied since the amplitude of the electric current remains small although the pressure and flow amplitudes are relatively large. Both flow and pressure saturation are definitely included in the orifice flow equations for both stages of the valve.

REFERENCES

1. See reference 3 of the paper.
2. See reference 1 of the paper.
3. See reference 2 of the paper.

The Lightweight Train—Its Power Supply for Auxiliaries

J. L. SWARNER
ASSOCIATE MEMBER AIEE

FOR MANY YEARS the "iron horse" rode alone having no equal or competition. The railroad conquered the wilderness with a belch of smoke and tied the far places together with ribbons of steel. Legends were woven around it and ballads created. Excesses brought about legal restrictions, but still the railroad prospered.

Born along with the present century was the horseless carriage which would not last, but did. The idea of mass production and the assembly line developed along with the automobile. The two teamed up to produce millions of offspring, rapidly and economically. The iron horse lost many of his riders to the family car.

During this so-called industrial revolution the railroads and their suppliers followed diametrically opposite paths from the general industrial pattern. Automobiles provided the method, but household appliances and practically every other industry employed mass production and assembly lines to provide a cheaper, better product to sell at a lower price.

Not so with the railroad industry, and by this term is included the equipment suppliers to the roads. Passenger cars were custom-built to individual railroad specifications. In the period 1945 through 1953, Pullman-Standard built 1,235 domestic nonsleeper passenger cars from 151 different floor plants for 41 different customers; 43% of these had less than three and 73.5% had less than seven cars per plan. In this same period, Pullman-Standard built 913 domestic sleeping cars, excluding troop sleepers, from 66 different floor plans and 36 different railroads. Of these, 21% had less than three and 42.5% had less than seven cars per plan. While there are six basic types of sleeping accommodations, 25 different types were engineered and built. To point this condition further, practically all of these sleeping cars were turned over to the Pullman Company to operate and to maintain. The Pullman Company, in order to operate this varied mass of cars, conducted extensive surveys and was able to achieve a degree of standardization by changing the cars after they took them over. A typical example concerns the batteries. Cars were received by the Pullman Company with 33 different types of batteries. With a minimum of structural change on the cars, this number was or is being reduced to nine as the batteries wear out.

With the appearance of the lightweight trains, built late in 1955 or in the early part of 1956, a 20-year cycle was completed. In 1934 Pullman-Standard, in co-operation with the Union Pacific Railroad, built a unit train with matched power, central auxiliary electric power, and light weight. It was a so-called tubular type, with a smaller cross section than standard. It weighed 1,110 pounds per foot of length, including the locomotive.

Since 1934, cars have been getting progressively heavier. Table I shows this trend from the 1,110 pounds per foot of the City of Salina to the 1,900 pounds per foot of length of the 1955 train.

The concept of something new in railroad passenger equipment was developed by necessity. A method must be found to make this operation desirable to the public and profitable to the railroads. There has been much publicity and speculation about the so-called lightweight trains. With their sleek lines, underslung bodies, matched power locomotives, and 120 miles per hour (mph) speed,

---

Paper 56-256, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE 1956 Winter General Meeting, New York, N. Y., January 30–February 3, 1956. Manuscript submitted November 21, 1955; made available for printing June 14, 1957.

J. L. SWARNER is with the Pullman-Standard Car Manufacturing Company, Chicago, Ill.

The author wishes to acknowledge information on the bus duct given by W. A. Kuhar and Gerald Panulla of the Roller-Smith Corporation, Bethlehem, Pa.

Table I. Train Weight Trend 1934 to 1955, Complete Trains Including Locomotive

Train	Length of Train, Feet-Inches	Weight of Train, Pounds	Weight per Feet, Pounds
City of Salina, 1934 (Union Pacific)	204- 5	232,540	1,110
Green Diamond (Illinois Central)	328- 6	476,800	1,146
City of Portland, 1936 (Union Pacific)	454-11	615,580	1,350
City of Los Angeles (Union Pacific)	713-11	1,005,640	1,410
City of Denver (Union Pacific)	864- 0	1,332,840	1,540
Conventional 6-car train, 1955	580- 3	1,104,300	1,900
Train *X*	497-10	383,000	770

Table II. Weight Comparison of Car Electric Systems

System	Per-Car Weight with Complete Electric Equipment, Pounds	Total Weight of Electric System for Train, Pounds
Axle-generator battery system, 110 volts d-c with amplidyne	12,586	75,516
Car equipment for train *X*, power source in locomotive	1,303	11,727
Central power equipment in locomotive		
New York Central	15,500	
New York, New Haven & Hartford	38,522	
Net difference for complete train system, axle generator versus train *X*		
New York Central		48,289
New York, New Haven & Hartford		25,267
Diesel-generator system 120 volts d-c	10,200	61,200
Net difference for complete train system, diesel generator versus train *X*		
New York Central		33,973
New York, New Haven & Hartford		10,951

Weights estimated: one locomotive for New York Central train, 300 kw; two locomotives for New York, New Haven & Hartford train, 250 kw in each locomotive.

they have been proclaimed as the salvation of railroad passenger business, in an era of increasingly tight competition; see New York Central train *X* in Fig. 1. But these trains will not live up to expectations, unless the following factors are given serious consideration and application:

1. Weight reduction without sacrifice of traditional railroad comfort and convenience to be accomplished by: use of new materials and techniques, providing energy for train auxiliaries from a common source. Of these two, the latter offers the greatest opportunity.
2. Sufficient standardization to permit production of car units in quantities, allowing for mass-production techniques.
3. A unit cost, which will allow the railroads to provide a better service, at equal or lower cost than other forms of transportation.
4. The reduction of maintenance and servicing, by intelligent design and use of proved components.

Using the foregoing as a basis, the electric system for lightweight trains can be selected.

Growth of Car-Auxiliary-Power Requirements

More than 125 years ago, the application of lighting started the demand for car auxiliaries for passenger comfort and safety. It was in 1825 that Thomas Dixon, operating a horsedrawn coach between Stockton and Darlington, on the Liverpool and Manchester Railroad, bought candles for his passengers to hold in the car. For the next 100 years, the primary motivation, for auxiliary energy on railroad passenger cars, was the need for light.

The invention of the electric incandescent lamp, preceded a few years by the development of the storage battery by Plante, opened the way for the application of electric auxiliary power to a car. In 1881, the London, Brighton and South Coast Railroad put the incandescent lamps and the storage battery together on a passenger car.

In 1894, the New Jersey Central Railroad made the first installation of an axle-generator system in the United States. This system gradually took over until 1934, when its use was practically universal.

Then came the application of air conditioning to passenger cars in the 1930's. The auxiliary-power demand jumped from 5 kw to 25 kw. The need for consideration of a more suitable power system was created. The recent building of cars with large glass areas, such as the dome-type cars and high-level cars, further increased car auxiliary loads due to air-conditioning requirements. Modern passenger cars demand anywhere from 20 to 40 kw of connected load.

Fig. 1. Artist's conception of New York Central's train X

Selection of Electric Power Source for a Train

There are three possible methods of providing power for auxiliaries on passenger cars. The axle-driven generator and associated battery, the diesel-driven generator for the individual cars, and a central source of power feeding the cars through a trainline. To apply any one of these to the lightweight train, the system must be compatible with the fundamental concepts of the new trains, as outlined previously.

Both the axle-generator system and the diesel-driven generator must be mounted under the car floor. The new trains, with their lowered floor heights, do not provide the necessary clearances.

In examining the characteristics of the axle-generator system further, the weights and cost must conform to the original premises. Table II is a comparison of a recent 110-volt d-c coach with an amplidyne invertor, a diesel-driven alternator system applied to the same type of car, and Pullman-Standard's train *X* which uses a central power source of 480 volts, 3 phase a-c. Fig. 2 is a similar comparison for a 32-volt car versus train *X*. Here again, the evident weight advantage of the central power source recommended its use.

Table III is an analysis of the horsepower requirements necessary for producing electric power by means of axle-driven generators. The efficiencies used and the horsepower to transport one ton of weight over level tangent track were taken from data determined in extensive Association American Railroads tests in 1937. This table indicates a 2.11-horsepower(hp)-per-kilowatt average from 30 to 90 mph. If axle-generated power was furnished for train *X*, it would require 50% of one locomotive's output,

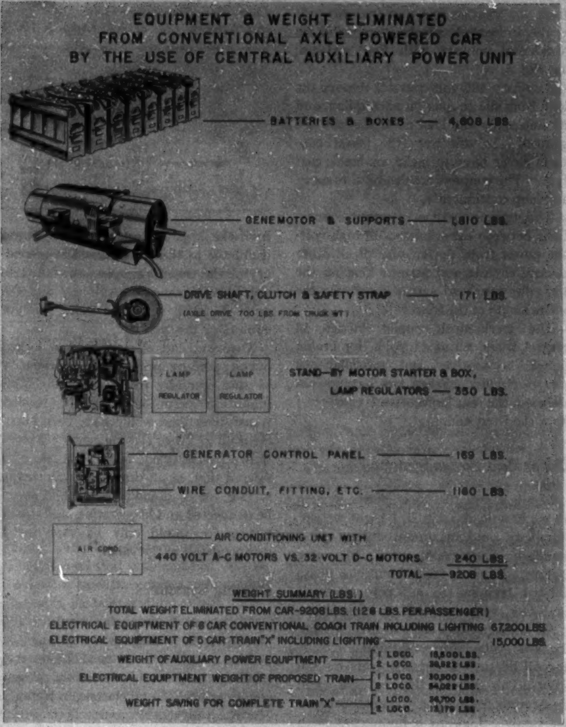

EQUIPMENT & WEIGHT ELIMINATED
FROM CONVENTIONAL AXLE POWERED CAR
BY THE USE OF CENTRAL AUXILIARY POWER UNIT
BATTERIES & BOXES — 4,608 LBS.
GENEMOTOR & SUPPORTS — 1,810 LBS.
DRIVE SHAFT, CLUTCH & SAFETY STRAP — 171 LBS.
(AXLE DRIVE 700 LBS. FROM TRUCK WT.)
LAMP REGULATOR
LAMP REGULATOR
STAND-BY MOTOR STARTER & BOX,
LAMP REGULATORS — 350 LBS.
GENERATOR CONTROL PANEL — 169 LBS.
WIRE CONDUIT, FITTING, ETC. — 1160 LBS.
AIR COND.
AIR CONDITIONING UNIT WITH
440 VOLT A-C MOTORS VS. 32 VOLT D-C MOTORS. 240 LBS.
TOTAL — 9208 LBS.
WEIGHT SUMMARY (LBS.)
TOTAL WEIGHT ELIMINATED FROM CAR-9208 LBS. (128 LBS. PER PASSENGER)
ELECTRICAL EQUIPTMENT OF 6 CAR CONVENTIONAL COACH TRAIN INCLUDING LIGHTING 67,200 LBS.
ELECTRICAL EQUIPTMENT OF 5 CAR TRAIN "X" INCLUDING LIGHTING — 15,000 LBS.
WEIGHT OF AUXILIARY POWER EQUIPTMENT — 1 LOCO. 15,500 LBS. 2 LOCO. 39,022 LBS.
ELECTRICAL EQUIPTMENT WEIGHT OF PROPOSED TRAIN — 1 LOCO. 30,500 LBS. 2 LOCO. 54,022 LBS.
WEIGHT SAVING FOR COMPLETE TRAIN "X" — 1 LOCO. 36,700 LBS. 2 LOCO. 13,178 LBS.

Fig. 2. Comparison: 32 volt d-c versus train X

Table III. Horsepower Required for Auxiliary Power

Six-Car Conventional Coach Train, Axle-Generator System and Train X

Item	30 MPH	50 MPH	70 MPH	90 MPH
Efficiency of generator and drive	79	77	75	72
Input hp at car axle for 20-kw general output	34	34.7	35.8	37.3
Input hp at rail at 5% loss	35.8	36.5	37.7	39.2
Drawbar hp required for six cars to generate 20-kw output	214.8	219.0	226.2	235.2
Drawbar hp required to transport each extra ton of equipment	0.24	0.53	1.06	2.16
Drawbar hp required to transport axle-generator equipment, one car	1.34	2.96	5.92	12.05
Drawbar hp required to transport six cars of axle-generator equipment	8.04	17.76	35.52	72.30
Total drawbar hp for six cars, sum of fourth and seventh items	222.12	236.76	261.52	307.5
Hp per kw output	1.77	1.95	2.18	2.56
Drawbar hp required to generate 270 kw for train *X* by means of axle generators	478	526	589	692
Engine hp required to generate 270 kw for train *X*, diesel-alternator set, 1.44 hp per kw	389 at all train speeds			
Drawbar hp required to transport train *X* auxiliary-power equipment				
One locomotive	1.86	4.10	8.21	16.74
Two locomotives	3.72	8.20	16.42	33.48
Total hp required for auxiliary power on train *X*				
One locomotive	390.86	393.1	397.21	405.74
Two locomotives	392.72	397.2	405.42	422.48

or 543-hp average, to turn the generators, whereas, only 397 hp would be necessary with a direct diesel-driven alternator. Since the primary source of power in both cases is a diesel engine, the advantage clearly lies with the central diesel-alternator set providing all train power.

Fig. 3 is a comparision of train resistances of the conventional 6-car coach train with axle-generator equipment and train *X*. Here is shown, graphically, the train resistance due to axle-generator drag. The primary purpose of a locomotive is to move the train. It appears that its horsepower could be more efficiently employed in increased acceleration rates and schedule speeds.

One important factor in obtaining the objective of low operating costs is the elimination of all items possible, and the arrival at a simple system made up of proven components, and available from mass-production sources. Referring to Fig. 2, there are six major items of equipment per car, which are special to the railroad industry and which require constant servicing and maintenance. If the axle generator and drive were replaced by a small diesel generator, the status is the same except for the increased function of diesel-engine repair, and fuel-oil and cooling service. All of this equipment, except for control panels, is mounted under the car where it is subjected to the worst possible operating conditions and where servicing and replacement are inconvenient and uncomfortable.

The car components of the electric power system for train *X* are as follows:

1. Locomotive Equipment
   a. Diesel generator.
   b. Engine-generator controls.
2. Car Equipment
   a. Three-phase bus-bar trainline.
   b. Automatic 3-phase trainline couplers.
   c. Power transformer.
   d. Magnetic contactors.

With one exception, these are items common to industry in general, and are produced by a number of reliable manufacturers at highly competitive prices. To illustrate the comparision between costs of equipment now used on passenger cars and what it is possible to use, Table IV is an example for d-c motors versus a-c motors in the horsepower ratings normal to railroad cars.

Selection of Power Trainline

From a previous study made in 1939, a trainline of two no. 4/0 conductors in parallel would be required for a 300-kw trainline. The weight of this trainline, for

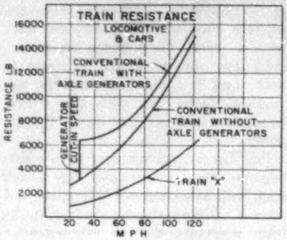


Fig. 3. Train resistance

train *X*, would have been 5,684 pounds, exclusive of end connectors or couplers.

An aluminum bus-bar system of standard manufacture, completely assembled with covers and hangers, rated at 600 amperes (amp), weighed 6 pounds per foot or 2,307 pounds for the train. The aluminum bus-bar system was selected. Voltage-drop calculations are shown in the Appendix.

The total maximum load on the trainline is 270 kw, or 325 amp at −20 F (degrees Fahrenheit) in the heating season. It is hoped that sufficient data can be obtained during the train's operation to arrive at a load factor for future trainline calculations.

Power Trainline Connectors

Train *X* is made up of four cars, each consisting of two permanently connected units, and a center car; see Fig. 4. The trainline is carried between the two permanently coupled ends of the 2-section car by means of bolted jumpers, near the roof of the car.

At the coupled ends, the power trainline is coupled automatically at the same time as the mechanical coupling. The power trainline coupler is mounted above the inner diaphragm, where it is protected from the weather and road-bed dirt. It is also so located as to be inaccessible for accidental contact with live surfaces; see Fig. 5.

Auxiliary 480-volt contacts remove the load from the coupler on uncoupling, and do not apply the load until the cars are coupled. In this way, the power contacts never have to make, or break current. The contacts are designed to carry 400 amp continuously.

This unique method of power connections between cars removes all high-voltage power from the vicinity of all other electric circuits and permits keeping the 480 volts overhead, out of reach, for the entire length of the train.

The mechanical coupler carries 42 control wires, all at 64 volts, for brake, signal, locomotive, and miscellaneous control circuits. When two locomotives are used, the rear locomotive is controlled from the front end.

Car or Unit Power System

Fig. 6 shows the basic power system on each unit of train *X*. Light weight, simplicity, and maximum utilization of standard components were governing factors. Power is taken from the 3-phase bus-bar trainline to a 3-pole magnetic contactor, with overloads on each phase.

All loads possible are fed directly from the 480 volts. Air-conditioning motors and all heating elements operate on 480 volts. This permits the use of a small transformer for lighting and miscellaneous loads only. Lighting is fluorescent, with 199/216-volt ballasts. The few incandescent lamps operate on 120 volts.

The transformer is connected delta-Y. The primary side is 480 volts delta-connected, the secondary is Y-connected for 208/120 volts.

On stand-by service, the insertion of the stand-by power plug isolates the unit from the train by locking out the power contactor to the trainline. The secondary of the transformer, i.e., the 208/120-volt side, becomes the primary and 480 volts is taken from the delta side for circuits of this voltage.

Table IV. Comparative Cost D-C Versus A-C Motors; Cost of A-C Motor Is 100%

Size, Hp	Cost of 220/440-Volt 3-Phase A-C Motor, Per Cent	Cost of 32-Volt D-C Motor, Per Cent
25	100	372
15	100	410
2	100	204
1	100	330
1/2	100	227

The overhead electric heat is locked out on stand-by, primarily due to capacity of yard stand-by power supplies. This limitation permitted using a smaller transformer than would have been possible if all 27 kw of electric heat were to be drawn from wayside power. It was considered that, at all times the train was in service, the head-end power system would be connected and in operation, and stand-by service would only be used for a detached unit in the yards.

Heating System

The same basic requirements for a modern train of the new design must be applied to all components. The heating system, traditionally steam on present equipment, will be all electric on train *X*. The following three systems were given serious consideration and engineering study:

1. A water-loop system, in which water is passed through a heat exchanger in the locomotive where it absorbs waste engine heat. This water is then circulated through a closed loop, the length of the train. Controls at each car select the amount of hot water necessary to maintain car tempera-

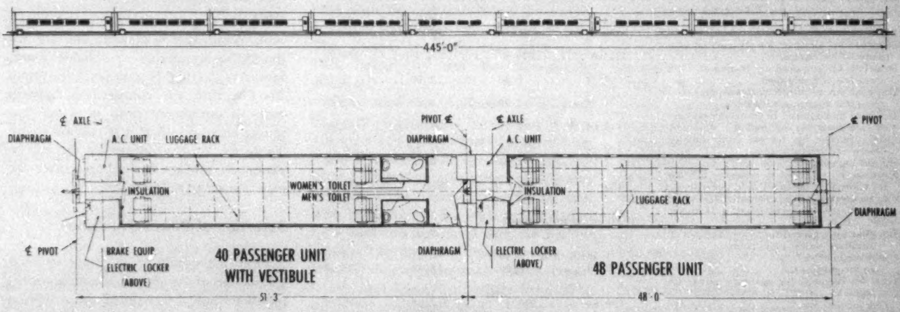


Fig. 4. Consist of train X

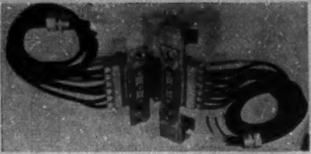


Fig. 5 (above). Power coupler

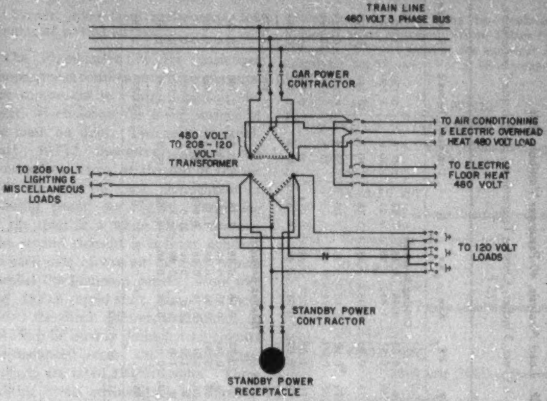

TRAIN LINE
480 VOLT 3 PHASE BUS
CAR POWER CONTRACTOR
TO AIR CONDITIONING & ELECTRIC OVERHEAD HEAT 480 VOLT LOAD
480 VOLT TO 208 - 120 VOLT TRANSFORMER
TO 208 VOLT LIGHTING & MISCELLANEOUS LOADS
TO ELECTRIC FLOOR HEAT 480 VOLT
N
TO 120 VOLT LOADS
STANDBY POWER CONTRACTOR
STANDBY POWER RECEPTACLE

Fig. 6 (right). Basic power circuit, train X

Table V. Comparative Weights of Heating System's Trailing Units

System	Weight Per Unit, Pounds	Weight Per Train, Pounds
Water loop	1,434	12,906
Steam	470	4,230
Unit boiler and radiation	936	8,424
Electric	150	1,350

ture. In warm weather, the water is chilled by a refrigeration system in the locomotive and is used to cool the cars.

2. A unit oil-fired boiler on each car, for either hot water or steam heating.
3. All-electric heating.

The water-loop-system engineering studies showed that sufficient heat would be available from the engines to maintain a desirable car temperature, except under light engine loads in severe weather. Electric heat was used to supplement when engine heat was insufficient. This system also indicated desirable operating economies by use of engine waste heat. However, other factors influenced against its use. These were: size of trainline pipe required, complexity of providing smooth flow, automatic coupling of the fluid line between units, antifreeze protection, and insulation of trainline and intercar connections for chilled water.

The unit oil-fired burner, one per car, required an electrical supply for the operating motor and controls, fuel tanks, filling devices, safety devices, and water supply. While this system of producing heat is more efficient than the generation of electric power by diesel engine, and conversion of power to heat, the maintenance of the oil-burner installation would more than offset any gain in efficiency. The filling of fuel tanks, establishment of wayside fueling facilities, and numerous units to maintain were not considered desirable.

The conventional steam-heating system, where steam is fed from a boiler in the locomotive through a steam trainline, is subject to relatively high losses in transmission, possible freeze-up, and complexity of providing automatic coupling between units. A steam boiler would have to be installed in the locomotive, in addition to the auxiliary generator for other auxiliaries.

The electric system offered the most over-all advantages; see Table V. Experience has proved the very low maintenance cost of this method of heating. Transmission losses are very low, and methods of coupling transmission lines between cars is simple, compared with other systems. Only one source of energy is required, per train, for all auxiliaries. Only one energy transmission line is necessary.

The electric heating system is composed of low-maintenance items of standard manufacture, namely strip heaters and contactors.

The electric heating system has two elements: the overhead unit, incorporated in the air-conditioning unit to temper fresh air, consists of an 18-kw finned-rod unit; the floor heat is made up of strip heaters, approximately 6 feet long, having a total capacity of 9 kw. Only two sizes of floor heat strips are used through-

Fig. 7. Air-conditioning units, air-intake side

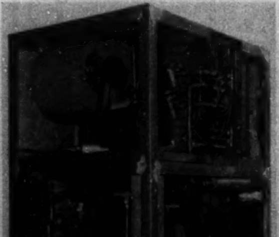


Fig. 8. Air-conditioning unit, cover removed

Fig. 9. Car interior

Table VI. Load Analysis, 48-Passenger Unit

Four Cars Per Train; Length, 48 Feet

Items	Kw	PF	Kva	Kv	I_L	ϕ, Degrees	$I_L \cos\phi$	$-jI_L \sin\phi$
Cooling Season*								
Compressor motor	5.60	0.85	6.58	0.48	7.92	31.8	6.730	4.17
Condensor motor	1.40	0.81	1.84	0.48	2.21	35.9	1.790	1.30
Fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct-diffused light	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Indirect light	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Miscellaneous light	0.45	1.0	0.45	0.48	0.541	0	0.541	0
Coffee warmer	0.60	1.0	0.60	0.48	0.721	0	0.721	0
	10.47						12.458	8.189
Heating Season†								
Fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct-diffused light	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Indirect light	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Miscellaneous light	0.45	1.0	0.45	0.48	0.541	0	0.541	0
Coffee warmer	0.60	1.0	0.60	0.48	0.721	0	0.721	0
Overhead heat Δ	18.0	1.0	18.0	0.48	21.6	0	21.600	0
Floor heat Δ	7.5	1.0	2.5	0.48	9.03	0	9.030	0
	28.88						34.568	1.419

* Without night light $I_L = 15.3/\underline{32.4^\circ}$, pf = 0.845.

† Pf = 0.998.

Table VII. Load Analysis, 40-Passenger Center Car

One Car Per Train; Length, 48 Feet; Outside Temperature, 20 F; Inside Temperature, 75 F

Items	Kw	PF	Kva	Kv	I_L	ϕ, Degrees	$I_L \cos\phi$	$-jI_L \sin\phi$
Cooling Season*								
Compressor motor	5.60	0.85	6.58	0.48	7.92	31.8	6.730	4.17
Condensor motor	1.49	0.81	1.84	0.48	2.21	35.9	1.790	1.30
Fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct-diffused light	0.99	0.90	1.10	0.48	1.32	25.8	1.190	0.575
Indirect light	0.99	0.90	1.10	0.48	1.32	25.8	1.190	0.575
Miscellaneous light	0.15	1.0	0.15	0.48	0.181	0	0.181	0
	9.97						11.847	7.517
Heating Season†								
Fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct-diffused light	0.99	0.90	1.10	0.48	1.32	25.8	1.190	0.575
Indirect light	0.99	0.90	1.10	0.48	1.32	25.8	1.190	0.575
Miscellaneous light	0.15	1.0	0.15	0.48	0.181	0	0.181	0
Overhead heat Δ	18.00	1.0	18.0	0.48	21.6	0	21.600	0
Floor heat Δ	9.00	1.0	3.0	0.48	10.83	0	10.830	0
	29.88						35.727	2.047

* Without night lights $I_L = 14/\underline{32.4^\circ}$, pf = 0.845.

† Pf = 0.997.

Table VIII. Load Analysis, 40-Passenger Vestibule Unit

Four Cars Per Train; Length, 51 Feet, 3 Inches; Outside Temperature, 20 F; Inside Temperature, 75 F

Items	Kw	PF	Kva	Kv	I_L	ϕ, Degrees	$I_L \cos\phi$	$-jI_L \sin\phi$
Cooling Season*								
Compressor motor	5.6	0.85	6.58	0.48	7.92	31.8	6.730	4.17
Condensor motor	1.49	0.81	1.84	0.48	2.21	35.9	1.790	1.30
Air-circulation fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct diffused lighting	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Indirect lighting	0.85	0.90	0.94	0.48	1.13	25.8	1.020	0.492
Miscellaneous lighting	0.76	1.0	0.76	0.48	0.914	0	0.914	0
Water cooler	0.19	0.60	0.32	0.48	0.385	53.2	0.231	0.308
	10.43						12.406	7.578
Heating Season†								
Air-circulation fan motor	0.75	0.76	0.98	0.48	1.18	40.5	0.766	0.897
Direct diffused light	0.79	0.90	0.88	0.48	1.06	25.8	0.955	0.411
Indirect light	0.85	0.90	0.94	0.48	1.13	25.8	1.020	0.492
Miscellaneous light	0.76	1.0	0.76	0.48	0.914	0	0.914	0
Overhead heat Δ	18.0	1.0	18.0	0.48	21.6	0	21.600	0
Floor heat Δ	9.0	1.0	3.0	0.48	10.83	0	10.830	0
Water cooler	0.19	0.60	0.32	0.48	0.385	53.2	0.231	0.308
	30.34						36.316	2.108

* Without night light $I_L = 14.5/\underline{31.5^\circ}$, pf = 0.853.

† Pf = 0.997.

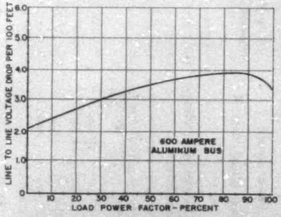


Fig. 10. Line-to-line voltage drop for train X

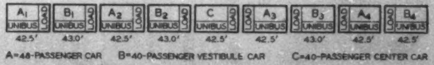


Fig. 11. Schematic of train X

out the train. The vestibule section, of the 2-section car, has a special-length heater in the toilets. All other units in the train are exactly alike and interchangeable.

To meet the need for two steps of heating, and to save weight and cost, the heaters are connected in delta, for full heat, and reconnected in Y for low heat. Therefore, there are 18 kw of electric heat at the high setting and 6 kw at the low connection, for the overhead air tempering. On the floor heat, the high setting is 9 kw, the low setting is 3 kw.

Air-Conditioning System

Each unit of train *X* has a 6-ton air-conditioning unit, which is a complete package, assembled, charged with Freon 12, and tested at the factory; see Figs. 7 and 8.

The unit is designed to be installed in the car through removable panels on the side of the car. The removable panels contain the condenser air inlet, the fresh-air grille, and the exhaust outlet. Electric connections to the units are plugs and receptacles. Duct connections must be completed, for car air supply and recirculated air, after the unit is bolted in place. Condenser air discharge is through the bottom of the car.

Here again, every effort was made to design an air-conditioning system which was lightweight, exactly identical for each car unit, easily replaceable, and made up of proven components The compressor is the sealed type. All motors are 440 volt, 3 phase, 60 cycle. All contactors and controls are mounted in a box, as an integral part of the unit. With spare units available, the railroad will be able to service them in the shop under ideal conditions, and not hold a car out of service.

In the design of the package system, it has been determined that any of the refrigeration equipment suppliers can produce a unit to be applied in the same location, and to conform to outside dimensions and air openings. In this way, standardization is achieved without limiting selection of equipment.

Lighting

The design of the lighting system is based on the same general stipulations as for the rest of the train. In addition, the illuminations will:

1. Eliminate glare sources.
2. Provide well-diffused illumination.
3. Maintain a high level of 40 foot-candles on the reading plane.
4. Permit group lamp replacement and cleaning at normal intervals.

The illumination in the passenger compartment consists of two components. One component is a large luminous element, which forms the lower surface of the main air duct. Two rows of rapid-start 48T12 fluorescent lamps are mounted on strips, at the top of the air duct, to provide even illumination of diffusing panel; see Fig. 9. The inside of the duct is a white plastic coating. The second element is indirect lighting, on each side of the air diffusers, which parallel the luminous panels. Here too, the 48T12 rapid-start lamp is used. Either the direct diffuser, the indirect, or both can be used as desired. All circuits are balanced across the 3-phase lines. Ballasts are rated 199/216 volts. Three lighting levels are available at approximately 15 foot-candles with indirect only, 30 foot-candles with direct diffused only, 40 to 50 foot-candles with both. The use of cool white deluxe lamps provides good color rendition to complement the color schemes.

An incandescent night-light circuit is included in the direct-diffused element, to provide a very low level, for service where passengers might desire to sleep.

Emergency lights are provided in each unit by means of a standard 6-volt storage-battery unit, complete with automatic charging facilities.

Conclusions

To arrive at a low first cost for a lightweight train with low operating cost, the power for train auxiliaries must come from a central source.

The use of equipment peculiar to the railroads, such as axle generators, battery, or small individual power plants, is unnecessary and undesirable.

The use of a nominal 440-volt 3-phase system allows the use of standard, proven, industrial equipment, available from many different sources at competitive prices.

The use of 3-phase power permits an air-conditioning system employing a sealed compressor and brushless motors for fan and condensor drives.

The heating system achieves the utmost in simplicity, with low maintenance.

Appendix

The load and trainline voltage-drop calculations for train *X* are shown in Tables VI through VIII. The end connections are not considered. Fig. 10 shows the line-to-line voltage drop for train *X*, and Fig. 11 is a schematic of train *X*. Using the calculated values of power factors (pf) from Tables VI through VIII, the current for each car is calculated as follows. For a 48-passenger unit

$$I_L = \frac{kw}{\sqrt{3}V_L \cos\theta} = \frac{9.97\times10^3}{(\sqrt{3})(480)(0.845)} =$$

14.2 amp (cooling season)

$$I_L = \frac{29.88\times10^3}{(\sqrt{3})(480)(0.997)} =$$

36.2 amp (heating season)

For a 40-passenger vestibule unit

$$I_L = \frac{10.43\times10^3}{(\sqrt{3})(480)(0.853)} =$$

14.7 amp (cooling season)

$$I_L = \frac{30.34\times10^3}{(\sqrt{3})(480)(0.997)} =$$

36.5 amp (heating season)

For a 40-passenger center unit

$$I_L = \frac{10.47\times10^3}{(\sqrt{3})(480)(0.845)} =$$

15.0 amp (cooling season)

$$I_L = \frac{28.8\times10^3}{(\sqrt{3})(480)(0.998)} =$$

34.8 amp (heating season)

In the cooling season, with pf = 0.85, the following may be calculated

voltage drop = 3.91 volts per 100 feet

For the 600-amp bus

voltage drop = 6.51×10^{-5} volts per amp per foot

The current I_L is calculated for a 48-passenger car (A), a 40-passenger vestibule car (B), and a 40-passenger center car (C) as follows

$I_L = 14.2$ amp at 0.85 pf for A
$I_L = 14.7$ amp at 0.85 pf for B
$I_L = 15.0$ amp at 0.85 pf for C

References

1. REALISTIC GOALS FOR RAILWAY PASSENGER CAR DESIGN, T. C. Gray. *Mechanical Engineering*, New York, N. Y., vol. 77, Mar. 1955, pp. 225–30. (condensed)

2. THE STEAM LOCOMOTIVE (book), Ralph P. Johnson. Simmons-Boardman Publishing Corporation, New York, N. Y., 1942, revised 1944.

3. CAR LIGHTING BY ELECTRICITY (book), Charles W. T. Stuart. *Ibid.*, 1923.

◆

Discussion

H. F. Brown (Gibbs & Hill, Inc., New York, N. Y.): This is a timely paper. Since so many different voltages have been used both in the United States and abroad for train lighting and heating, it would be of interest to learn from the author just what systems were considered and compared, economically, before the 440-volt 3-phase 60-cycle system was finally adopted.

It may well be that such a study, if made, would have an important bearing later, if railroad electrification with commercial frequency were to be adopted in this country. The 1,000-volt single-phase a-c trainline with simple single-conductor jumpers between cars would seem to have attractive possibilities. This is widely used in Europe.

In Fig. 3 is shown a curve of train resistance for train *X* compared with a conventional train. If the values shown for the train resistance for train *X* are converted into horsepower per ton, the values obtained by so doing would be as follows:

Speed in mph	30	50	70	90	120
Hp per ton	0.48	1.32	2.5	4.5	9.6

The values shown for horsepower per ton in Table III, the fifth item, are much lower than these values. Since these figures have an important bearing on the further calculations tabulated in Table III, it would appear to be in order to re-examine the values shown in this item.

In the eleventh item of Table III, the statement is made that the engine horsepower requirements are constant throughout the speed range when taken from a head-end supply. In one sense this is true, but since the prime mover must also move the weight of the auxiliary-power generating equipment, whether located on the cars or on the locomotive, it must devote the same increase in horsepower for each ton of added weight in either case, for a fair comparison, as the speed of the train increases.

Actually, the locomotive must supply nearly 30 hp more to move the weight of the head end and car equipment shown for train *X* at 90 mph than for 30 mph. If this power for auxiliaries comes from the main engine, there is this much less for traction. It is not much, but it should be taken into account for a fair comparison with axle-generator equipment. Train *X* is shown as weighing 770 pounds per foot including the locomotive. On this basis, it is only fair to include in the paper the horsepower of the engine for traction purposes since this will greatly influence the permissable maximum locomotive weight.

Since the lightweight train has been associated with high-speed trains in all recent publicity in the popular and technical press, attention is called to the increasing amount of horsepower per ton required for higher speeds than those now attained by conventional trains. To obtain speeds of 100 mph, or more, there must be available for traction capacities greater than 10 horsepower per ton to accelerate a train on level track up to these high speeds. Since self-propelled locomotives of any type now made in the United States at least are limited to about 13 or 14 hp per ton of locomotive weight, and since this is rapidly reduced as train weight, no matter how light, is added, it will be seen that it is almost a physical impossibility to design even a lightweight train for such speeds as 120 mph, unless the prime mover is designed to carry passengers.

In Europe, very high speed tests were recently made by the French railways, three cars weighing 101 tons were hauled by an electric locomotive weighing 80 tons. When this train of 181-ton total weight attained the maximum speed of 206 mph, it was taking more than 10,000 hp from the overhead contact system. This is just an indication of the way the power requirements increase as the speed increases, which seems to have been overlooked in some of the so-called high-speed-train designs being proposed in the United States.

The lightweight train will certainly save power at all speeds, when compared with conventional trains at the same speeds. The real problem is to obtain enough power for the speeds desired. It would appear that electrification offers the only practical solution for high speeds at the present time.

Multiple-Unit-Rectifier Motive Power—Inductive Co-ordination Considerations on the New York, New Haven & Hartford Railroad

L. J. HIBBARD
MEMBER AIEE

F. T. GARRY
ASSOCIATE MEMBER AIEE

G. N. LOOMIS
NONMEMBER AIEE

WHEN the New York, New Haven & Hartford Railroad ordered ten ignitron locomotives and 100 rectifier multiple-unit (MU) cars for use in their electrified territory, the communications engineers of the New Haven Railroad and of the Southern New England Telephone Company (SNET) were concerned with the possible increased levels of noise that might be encountered on their exposed lines.

While it was known from Pennsylvania Railroad tests and Westinghouse laboratory tests that filters could be applied to reduce the $I \cdot T$* product of such equipment, a question arose as to the cumulative effect of a combination of ignitron locomotives and MU cars operating in the same vicinity.

A pioneer rectifier MU car has been in continuous service with originally applied apparatus on the Pennsylvania Railroad for almost 6 years, since July 14, 1949, without an a-c filter. However, this car only draws approximately 52 amp (amperes) from the 11-kv catenary during acceleration.[1]

Two pioneer 6,000-hp (horsepower) rectifier locomotives have been in use on the Pennsylvania Railroad since July 1951. Each 2-cab locomotive normally draws 750 to 800 amp from the 11-kv trolley under full-load conditions, but sometimes draws 800 to 850 amp for 15 to 20 minutes with occasional short-time peaks of 1,000 amp.

There was no evidence that the two 6,000-hp locomotives had a directly additive effect on each other and laboratory tests definitely indicated that the additive effect of a large number of rectifier motive-power units should never be very pronounced. However, the New Haven commitment involved the first large-scale use of rectifier motive power in the world and test results from the application have been awaited with keen interest.

The 100 rectifier MU cars represent a potential of approximately 40,000-wheel-hp continuous load and a potential of approximately 62,000-wheel-hp accelerating load. The ten rectifier locomotives represent a potential of approximately 40,000-wheel-hp continuous and approximately 80,000-wheel-hp accelerating load.

---

*$I \cdot T$ represents the current in amperes supplied to the motive-power unit times the telephone influence factor which has a different weighting for each frequency in the voice range, based on telephone receiver response.

Paper 55-207, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Montreal, Que., Canada, June 24–28, 1957. This paper was previously presented as a Conference Paper at the 1955 Winter General Meeting. Manuscript submitted October 21, 1954; made available for printing May 17, 1957.

L. J. HIBBARD is with Westinghouse Electric Corporation, East Pittsburgh, Pa.; F. T. GARRY is with the Southern New England Telephone Company, New Haven, Conn.; and G. N. LOOMIS is with the New York, New Haven & Hartford Railroad Company, New Haven, Conn.

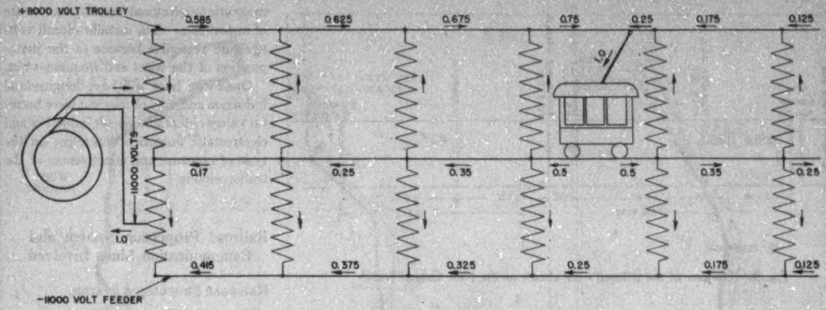

+11000 VOLT TROLLEY
0.585
0.625
0.675
0.75
0.25
0.175
0.125
1.0
11000 VOLTS
0.17
0.25
0.35
0.5
0.5
0.35
0.25
1.0
0.415
0.375
0.325
0.25
0.175
0.125
-11000 VOLT FEEDER

Fig. 1. Single-phase system with autotransformers

Factor Involved in Co-ordination Problem

Telephone circuits that pass through alternating magnetic fields will have extraneous currents induced in all connected telephone receivers. These extraneous currents interfere with telephonic transmission if they are in the audio- or noise-frequency range, and are of appreciable magnitude compared with normal voice currents.

The over-all effect of a given inductive exposure depends on three factors: 1. inductive susceptiveness of the telephone system, 2. inductive coupling, and 3. inductive influence of the power system.

If any of these three factors could be reduced to zero, the inductive effects would disappear. Inductive susceptiveness and inductive influence can be kept to a low value but cannot at the present state of the art be reduced to zero. To reduce couplings to zero would require large separations between the power and telephone systems, which is almost always impracticable in connection with power and telephone distribution systems and often difficult in connection with power transmission systems. The practical solution to most inductive problems, therefore, lies in keeping both susceptiveness and influence at a reasonably low value and using such measures as can be readily applied to limit coupling, e.g., telephone cable sheath shielding.

Control of telephone susceptiveness is largely a matter of controlling unbalance. Control of power-system influence is almost entirely a matter of controlling harmonics.

The voltage induced in telephone circuits can be resolved into two components:

1. Metallic-circuit induction. Unequal voltages will be induced in the two con-

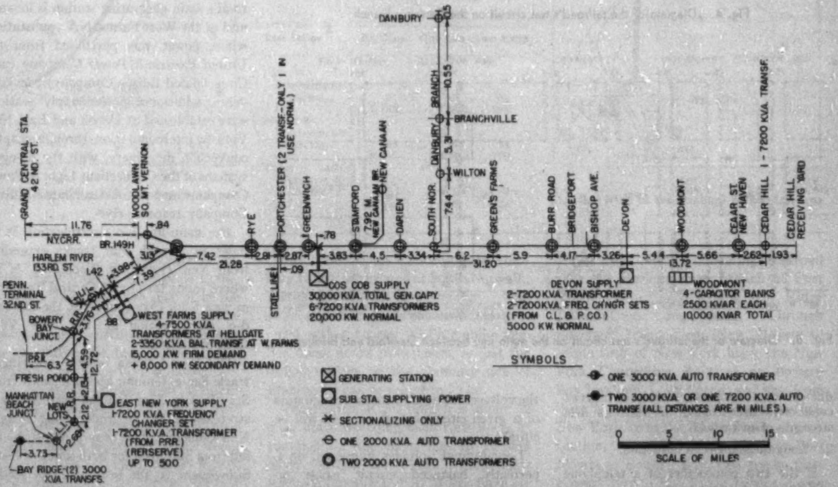

GRAND CENTRAL STA. 42 ND. ST.
WOODLAWN SO. MT. VERNON
RYE
PORTCHESTER (2 TRANSF.—ONLY 1 IN USE NORM.)
GREENWICH
STAMFORD
7.92 M. NEW CANAAN BR.
NEW CANAAN
DARIEN
SOUTH NOR.
DANBURY
10.5
10.55
BRANCHVILLE
DANBURY BRANCH
5.31
WILTON
7.44
GREEN'S FARMS
BURR ROAD
BRIDGEPORT
BISHOP AVE.
DEVON
WOODMONT
CEAAR ST. NEW HAVEN
CEDAR HILL 1 - 7200 KVA. TRANSF.
CEDAR HILL RECEIVING YARD
11.76
N.Y.C.R.R.
.84
BR.149H
3.13
7.42
21.28
2.81
2.87
.09
.78
3.83
4.5
3.34
6.2
31.20
5.9
4.17
3.26
5.44
13.72
5.66
2.62
1.93
HARLEM RIVER 133RD. ST.
1.42
3.98
7.39
PENN. TERMINAL 32ND. ST.
BOWERY BAY JUNCT.
.88
P.R.R.
6.3
N.Y.
3.76
4.59
12.72
FRESH POND
2.15
MANHATTAN BEACH JUNCT.
NEW LOTS
2.12
L.I.
2.68
3.73
BAY RIDGE-(2) 3000 KVA. TRANSFS.
STATE LINE
COS COB SUPPLY
30,000 KVA. TOTAL GEN. CAPY.
6-7200 KVA. TRANSFORMERS
20,000 KW. NORMAL
WEST FARMS SUPPLY
4-7500 KVA.
TRANSFORMERS AT HELLGATE
2-3350 KVA. BAL. TRANSF. AT W. FARMS
15,000 KW. FIRM DEMAND
+ 8,000 KW. SECONDARY DEMAND
EAST NEW YORK SUPPLY
1-7200 KVA. FREQUENCY CHANGER SET
1-7200 KVA. TRANSFORMER
(FROM P.R.R.)
(RERSERVE)
UP TO 500
DEVON SUPPLY
2-7200 KVA. TRANSFORMER
2-7200 KVA. FREQ. CH'NG'R SETS
(FROM C.L. & P. CO.)
5000 KW. NORMAL
WOODMONT
4-CAPACITOR BANKS
2500 KVAR EACH
10,000 KVAR TOTAL
SYMBOLS
GENERATING STATION
SUB. STA. SUPPLYING POWER
SECTIONALIZING ONLY
ONE 2000 KVA. AUTO TRANSFORMER
TWO 2000 KVA. AUTO TRANSFORMERS
ONE 3000 KVA. AUTO TRANSFORMER
TWO 3000 KVA. OR ONE 7200 KVA. AUTO TRANSF. (ALL DISTANCES ARE IN MILES)
0
5
10
15
SCALE OF MILES

Fig. 2. New York, New Haven & Hartford Railroad electrified zone showing sources and distribution of power, transformer locations, and spacing

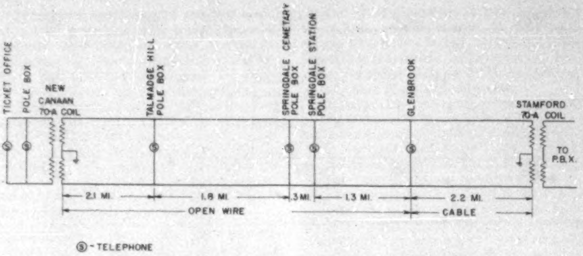

TICKET OFFICE
POLE BOX
NEW CANAAN 70-A COIL
TALMADGE HILL POLE BOX
SPRINGDALE CEMETARY POLE 90X
SPRINGDALE STATION POLE BOX
GLENBROOK
STAMFORD 70-A COIL
TO P.B.X.
2.1 MI.
1.8 MI.
.3 MI.
1.3 MI.
2.2 MI.
OPEN WIRE
CABLE
Ⓢ - TELEPHONE

Fig. 3. Diagram of the railroad's test circuit on the New Canaan branch

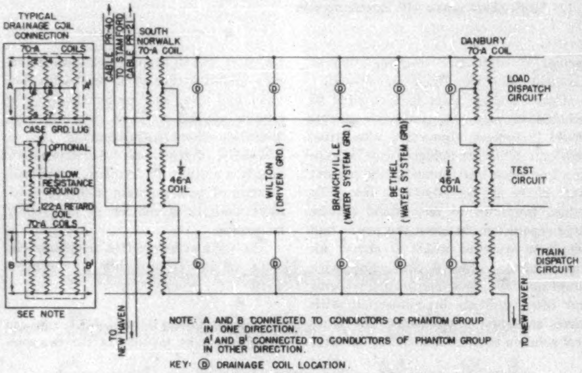

TYPICAL DRAINAGE COIL CONNECTION
70-A COILS
A
A'
CASE GRD. LUG
OPTIONAL
LOW RESISTANCE GROUND
122-A RETARD COIL
70-A COILS
B
B'
SEE NOTE
CABLE PR-40
TO STAMFORD
CABLE PR-21
SOUTH NORWALK 70-A COIL
46-A COIL
WILTON (DRIVEN GRD)
BRANCHVILLE (WATER SYSTEM GRD)
BETHEL (WATER SYSTEM GRD)
46-A COIL
DANBURY 70-A COIL
LOAD DISPATCH CIRCUIT
TEST CIRCUIT
TRAIN DISPATCH CIRCUIT
TO NEW HAVEN
TO NEW HAVEN
NOTE: A AND B CONNECTED TO CONDUCTORS OF PHANTOM GROUP IN ONE DIRECTION.
A' AND B' CONNECTED TO CONDUCTORS OF PHANTOM GROUP IN OTHER DIRECTION.
KEY: Ⓓ DRAINAGE COIL LOCATION.

Fig. 4. Diagram of the railroad's test circuit on the Danbury branch

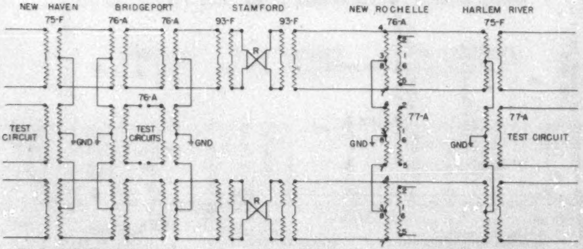

NEW HAVEN
75-F
BRIDGEPORT
76-A
76-A
STAMFORD
93-F
93-F
NEW ROCHELLE
76-A
HARLEM RIVER
75-F
R
76-A
TEST CIRCUIT
GND
TEST CIRCUITS
GND
77-A
GND
77-A
GND
TEST CIRCUIT
R
KEY R= VOICE AND RINGING REPEATER
NOTE: TEST CIRCUIT 16 PHANTOM ON QUAD OF 16-GA CABLE CONDUCTORS LOADED M-172-63

Fig. 5. Diagram of the railroad's test circuit on the main line between Stamford and Bridgeport

ductors of a telephone circuit if the two conductors are located in different field strengths of an a-c field.

2. Longitudinal-circuit induction.

If the two conductors of a telephone circuit are subjected to the same average field strengths by means of transpositions, the voltages induced in the two conductors of a given circuit by an a-c field will be equal.

Longitudinal voltages impressed on a perfectly balanced circuit cause no current in connected telephone receivers.

For telephone cable circuits, only the magnetic longitudinal coupling factors are of importance. The metallic-circuit voltages are negligible because of the juxtaposition of the wires and frequent twist.

Open-wire lines will have longitudinal induction and may or may not have harmful values of metallic-circuit induction and electrostatic induction depending on the type of exposure and effectiveness of the transpositions.

Railroad Propulsion System and Communication Lines Involved

Railroad Propulsion System

The electrification of the New York, New Haven & Hartford Railroad includes the main line from Woodlawn to New Haven, the New Canaan branch, the Danbury branch, the Harlem River branch, and the New York Connecting Railroad.

Power is supplied from overhead catenaries at 11,000 volts, 25 cycles, single phase. In the major part of the electrification, the trolley-rail circuit is fed by autotransformers, which, with the trolley rails and feeders, form a 3-wire 22,000/11,000-volt system having 22,000 volts between the trolleys and feeders, and 11,000 volts between rails and either trolleys or feeders. At the time of the completion of the Harlem River—New Haven communication cable in 1918, power was fed to the propulsion system only at Cos Cob, Conn., where the railroad's main generating station is located, and at the West Farms, N.Y., substation, where power was purchased from the United Electric & Power Company (now Consolidated Edison Company). In later years, additional power-supply stations were established at Devon and East New York by interconnection, through suitable conversion machinery, with the 60-cycle systems of the Connecticut Light & Power Company and the Consolidated Edison Company respectively.

The main-line a-c electrification is 64 miles long, with four tracks extending from New Haven to Woodlawn, N.Y., where trains for Grand Central Terminal in New York City transfer to the 650-volt d-c electrification of the New York Central Railroad.

The New Canaan branch is a single-track line extending about 8 miles from Stamford to New Canaan with frequent stretches of 1% upgrade, going toward New Canaan. This branch was stub-end fed from Stamford for both series of tests. Subsequent to the tests, a 22-kv feeder has been extended from Stamford to a transformer at New Canaan.

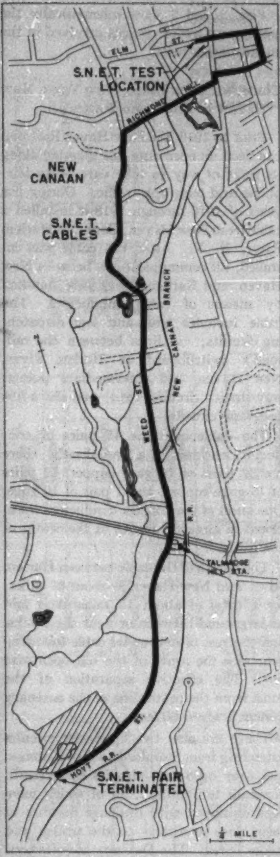

ELM ST.
S.N.E.T. TEST LOCATION
RICHMOND HILL
NEW CANAAN
S.N.E.T. CABLES
WEED ST.
NEW CANAAN BRANCH R.R.
TALMADGE HILL STA.
HOYT ST.
S.N.E.T. PAIR TERMINATED
½ MILE

Fig. 6. Map of the New Canaan branch test area

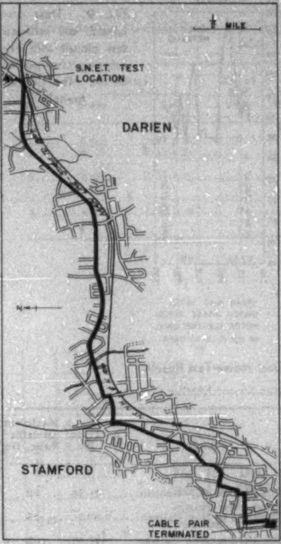

½ MILE
S.N.E.T. TEST LOCATION
DARIEN
N
STAMFORD
CABLE PAIR TERMINATED

Fig. 7. Map of the test area on the main line between Stamford and Bridgeport

Table I. Maximum I·T Product, Single MU Car

	One Car		
Test Series	No Cans	One Can	Two Cans
First			
Maximum db reading	60	60	52
I·T	1,728	1,728	692
Second			
Maximum db reading	63	60	54
I·T	2,440	1,728	776
Third			
Maximum db reading	63	59.5	53
I·T	2,440	1,640	776

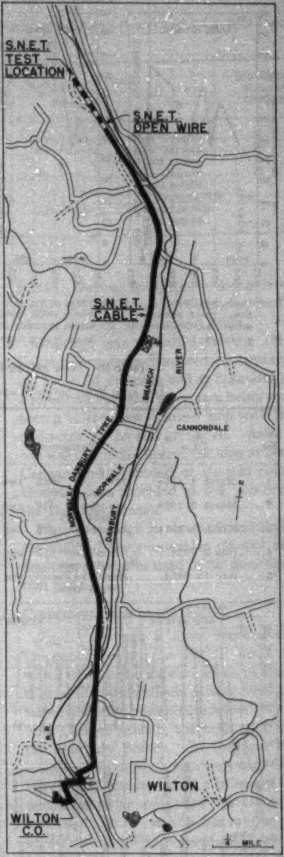

S.N.E.T. TEST LOCATION
S.N.E.T. OPEN WIRE
S.N.E.T. CABLE
BRANCH
RIVER
CANNONDALE
NORWALK DANBURY TPKE
NORWALK
DANBURY
N
R.R.
WILTON
WILTON C.O.
½ MILE

Fig. 8. Map of the Danbury branch test area

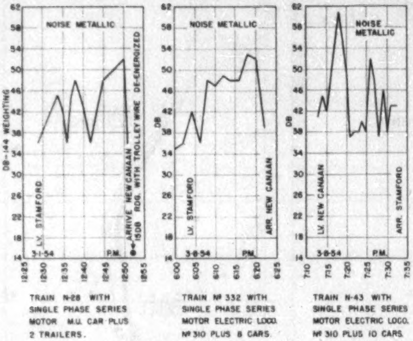
An installation of one series-track booster transformer at Springdale Cemetery and a 3-ohm air-core reactor in the trolley at Stamford was used to limit the magnitude and path of short-circuit currents during the tests. The booster transformer has since been removed.

The Danbury branch, also a single-track line, extends about 24 miles from South Norwalk to Danbury. High-speed circuit breakers are installed in the feeder and trolley wires on the Danbury side of the autotransformer at South Norwalk to limit the duration of short-circuit currents.

Autotransformers for feeding the Danbury trolley are located at South Norwalk, Wilton, Branchville, and Danbury. The autotransformer at Branchville is provided with a booster winding and a section break in the trolley wire, so that the trolley voltage on the Danbury side of Branchville can be raised to 12 kv.

The Harlem River branch is about 12 miles long, extending from New Rochelle Junction on the main line to Harlem River yard in New York. From New Rochelle Junction to Port Morris there are four tracks, and from Port Morris to Harlem River there are two.

The New York Connecting Railroad joins the Harlem River branch at Port Morris and extends to the Bay Ridge Terminal on Long Island. Through passenger trains from New Haven to points west of New York leave the New York Connecting Railroad at Bowery Bay Junction, Long Island, and pass from there through the Pennsylvania Station, New York, over the Pennsylvania electrification. There are four tracks from Port Morris to Bowery Bay, two tracks from Bowery Bay to Bay Ridge Terminal, and two tracks from Bowery Bay to Harold Avenue.



TRAIN N-28 WITH SINGLE PHASE SERIES MOTOR M.U. CAR PLUS 2 TRAILERS.

TRAIN Nº 332 WITH SINGLE PHASE SERIES MOTOR ELECTRIC LOCO. Nº 310 PLUS 8 CARS.

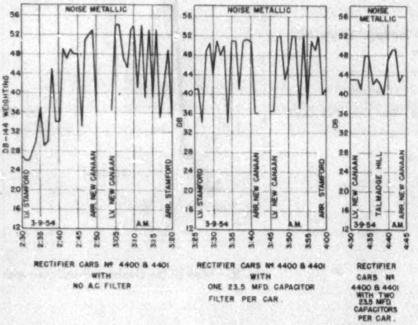
TRAIN N-43 WITH SINGLE PHASE SERIES MOTOR ELECTRIC LOCO. Nº 310 PLUS 10 CARS.

Fig. 9. Data obtained on railroad test circuit with existing motive power on the New Canaan branch

Table II. New Haven Railroad Noise-Test Results

Test Runs With Single-Phase Series-Motor Motive Power

Test No.	Date	Equipment	Route	Train No.	Maximum Metallic Noise, Db
1	March 1, 1954	MU car with two trailers	Stamford to New Canaan	N-28	52
2	March 8, 1954	locomotive 310 and eight cars	Stamford to New Canaan	332	53
3	March 8, 1954	locomotive 310 and ten cars	Stamford to New Canaan	N-43	61
4	July 8 1954	locomotive 315 and eight cars	South Norwalk to Danbury	140	50
5	July 13, 1954	electric locomotive and eight cars	South Norwalk to Danbury	356	50



RECTIFIER CARS Nº 4400 & 4401 WITH NO A.C. FILTER

RECTIFIER CARS Nº 4400 & 4401 WITH ONE 23.5 MFD. CAPACITOR FILTER PER CAR.

RECTIFIER CARS Nº 4400 & 4401 WITH TWO 23.5 MFD. CAPACITORS PER CAR.

Fig. 10. Data obtained on railroad test circuit with rectifier cars on the New Canaan branch

Table III. New Haven Railroad Noise-Test Results

Test Runs on MU Cars, March 9, 1954, on New Canaan Branch

Test No.	Time, A.M.		No. of MU Cars	Filters Per Car	Train		Maximum Metallic Noise on New Canaan Freight and Ticket Line, Db
	From	To			From	To	
1	2:18	2:52	2	0	Stamford	New Canaan	53
2	3:04	3:24	2	0	New Canaan	Glenbrook	54
3	3:25	3:42	2	1	Glenbrook	New Canaan	51.5
4	3:45	4:04	1	1	New Canaan	Glenbrook	52
5	4:30	4:44	2	2	New Canaan	Talmadge Hill and return	49
6	4:46	5:00	1	2	New Canaan	Talmadge Hill and return	51
	5:10		1*	2	New Canaan	Stamford	52

* Second car towed.

Figs. 1 and 2 show schematically the trolley and feeder circuits involved in the co-ordination tests.

Communication Lines, New York, New Haven & Hartford Railroad

Prior to 1919, the New Haven Railroad had used an open-wire line situated along its right of way to take care of its telephone and telegraph traffic. During the years of 1916 through 1918 it installed a cable from New Haven, Conn., to Harlem River, N. Y. After the cable was installed, all communications between New Haven and Harlem River were handled by means of cable conductors. The cable includes train-and load-dispatching circuits; tie lines between the railroad's switchboard at Harlem River, New Haven, and intermediate points; way-station circuits, etc.; and also a few telegraph circuits.

The cable contains 45 pairs of conductors enclosed in a lead sheath; there are 32 pairs of 16 gage copper, 12 pairs of 13 gage copper, and 1 pair of 10 gage. One quad of the 16 gage conductors is enclosed in a second sheath at the center of the cable.

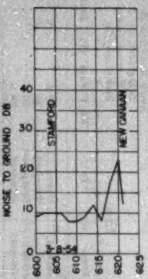
The length of the cable between Harlem River and New Haven is about 67 miles; for a total of about 15 miles it is laid underground between or near the tracks, and the rest of it is aerial cable following generally the route of the old open-wire line. The effective separation of the cable from the center line of the catenary system is about 60 feet.

There are also two open-wire circuits extending from Stamford to New Canaan, and four open-wire circuits from South Norwalk to Danbury, all of which have been equipped with drainage because of their close proximity to the trolley and feeder wires. The Danbury circuits consist of two physical and one phantom group of no. 9 Brown & Sharpe gage copper wire. Both the Danbury and New Canaan communication circuits are carried along the railroad right of way and are horizontally separated from the trolley by about 25 feet.

The New Haven Railroad operates circuits in a cable between Harlem River and Bay Ridge, carried for most of the distance along the right of way of the New York Connecting Railroad.

At points where the railroad circuits connect with those of the Bell System, repeating coils are used to separate the two metallic circuits.

Western Electric repeating coils are installed on the main-line circuits for drainage at the following locations: New



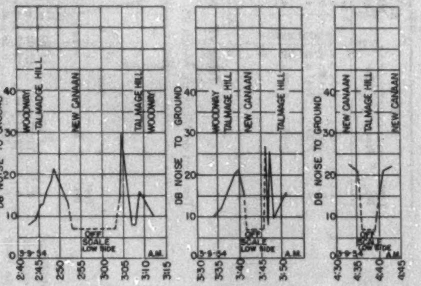
* TRAIN NO.332 WITH SINGLE PHASE SERIES MOTOR ELEC. LOCO. NO.30 PLUS 8 CARS

Fig. 11. Data obtained on SNET test circuit with existing motive power on the New Canaan branch

Table IV. SNET Noise-Test Results

Test Runs on New Canaan Branch with Single-Phase Series Motive Power and 2-Car MU Train, March 8 and 9, 1954

Test No.	Time		No. of MU Cars	Filters Per Car	Train		Maximum Noise to Ground
	From	To			From	To	
1	6:02	6:21 p.m.	0	0	Stamford	New Canaan	23
2	2:43	3:12 a.m.	2	0	Woodway	New Canaan and return	30
3	3:34	3:51 a.m.	2	1	Woodway	New Canaan and return, Talmadge Hill	27
4	4:33	4:43 a.m.	2	2	New Canaan	Talmadge Hill and return	22.5



RECTIFIER CARS NO. 4400 & 4401 WITH NO A.C. FILTER

RECTIFIER CARS NO.4400 & 4401 WITH ONE 23.5 MFD. CAPACITOR FILTER PER CAR

RECTIFIER CARS 4400 & 4401 WITH TWO 23.5 MFD. CAPACITORS PER CAR

Fig. 12. Data obtained on SNET test circuit with rectifier cars on the New Canaan branch

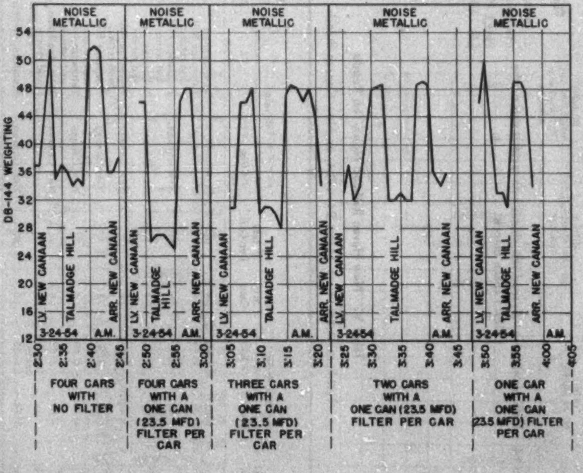


Fig. 13. Data obtained on railroad test circuit with rectifier cars on the New Canaan branch

Haven, Devon, Bridgeport, South Norwalk, Port Chester, New Rochelle, South Mount Vernon, and Harlem River.

Drainage was provided for the New Canaan loop from Stamford by connecting the *70-A* drainage coil at Stamford as a repeating coil, the secondary being connected to the New Canaan loop with the mid-point of the coil grounded. A *70-A* coil was connected across the loop at New Canaan and grounded at its mid-point. Drainage was provided at five points on the Danbury branch.

Figs. 3 through 5 show schematically the railroad communication circuits and drainage provisions used in the co-ordination tests.

COMMUNICATION LINES, SNET

Figs. 6 through 8 show, respectively, the exposures to the SNET plant on which measurements were made of longitudinal-induction effects arising from the railway motive-power equipment on the New Canaan branch, the main line, and the Danbury branch. In the first two instances, all cable circuits were used with the far-end terminated in a balanced impendance to ground. On the Danbury branch, the test location was at the end of an open-wire line about 1/2 mile in length. These open-wire circuits were connected to cable pairs that had been terminated in a balanced impedance to ground at the Wilton central office.

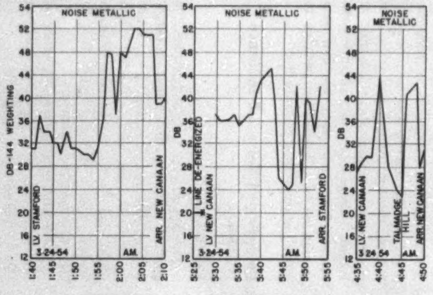
Co-ordinated Noise and Influence Tests

EXPOSURE LOCATIONS

The first series of tests to determine the effect of rectifier motive power on neighboring communication circuits was made on March 9, 1954, with two rectifier cars (4400 and 4401) on the New Canaan branch, with a stub-end trolley feed to New Canaan from Stamford.

The second series of tests was made on March 24, 1954, with six rectifier cars on

(*Continued on page 424.*)

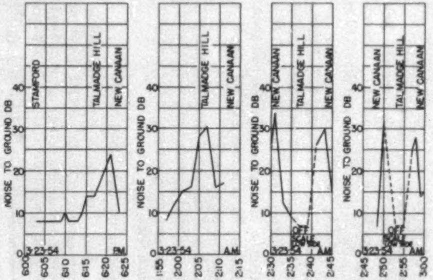


SIX RECTIFIER MU. CARS WITH NO A.C. FILTER.

SIX RECTIFIER MU. CARS WITH A TWO CAN (47 MFD.) FILTER PER CAR.

FOUR RECTIFIER MU. CARS WITH A TWO CAN FILTER (47 MFD) PER CAR.

Fig. 14. Data obtained on railroad test circuit with rectifier cars on the New Canaan branch



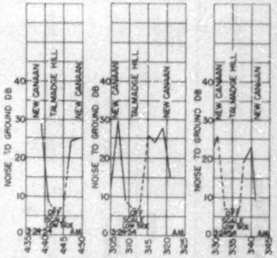
A.C. TRAIN WITH 6 CARS

MU. TRAIN WITH 6 CARS & NO FILTERS

MU. TRAIN WITH 4 CARS AND NO FILTERS

MU. TRAIN WITH 4 CARS & ONE FILTER PER CAR

Fig. 15. Data obtained on SNET test circuit with rectifier cars and an a-c train on the New Canaan branch



MU. TRAIN WITH 4 CARS & TWO FILTERS PER CAR

MU. TRAIN WITH 3 CARS & ONE FILTER PER CAR

MU. TRAIN WITH 2 CARS & ONE FILTER PER CAR

Fig. 16. Data obtained on SNET test circuit with rectifier cars on the New Canaan branch

Table V. New Haven Railroad Noise-Test Results

Test Runs on MU Cars, March 24, 1954, on New Canaan Branch

Test No.	Time, A.M. From	Time, A.M. To	No. of MU Cars	Filters Per Car	Train From	Train To	Maximum Metallic Noise on New Canaan Freight and Ticket Line, Db
1	1:33	2:10	6	0	Stamford	New Canaan	52
2	2:30	2:45	4	0	New Canaan	Talmadge Hill and return	52
3	2:49	3:00	4	1	New Canaan	Talmadge Hill and return	48
4	3:05	3:21	3	1	New Canaan	Talmadge Hill and return	48.5
5	3:25	3:43	2	1	New Canaan	Talmadge Hill and return	49
6	3:49	3:58	1	1	New Canaan	Talmadge Hill and return	50
7	4:35	4:50	4	2	New Canaan	Talmadge Hill and return	44
8	4:59	5:23 (reactor cut out)	6	2	New Canaan	Talmadge Hill and return	48
9	5:28	5:53	6	2	New Canaan	Stamford	45

Table VI. SNET Noise-Test Results

Test Runs on New Canaan Branch With Single-Series Motive Power and MU Cars, March 23 and 24, 1954

Test No.	Time From	Time To	No. of MU Cars	Filters Per Car	Train From	Train To	Maximum Noise to Ground
1	6:03	6:24 p.m.	0	0	Stamford	New Canaan	24
2	1:52	2:13 a.m.	6	0	Stamford	New Canaan	30.5
3	2:29	2:45 a.m.	4	0	New Canaan	Talmadge Hill and return	34
4	2:48	3:00 a.m.	4	1	New Canaan	Talmadge Hill and return	31.5

Table VII. New Haven Railroad Noise-Test Results

Test Runs on MU Cars, March 26, 1954, on Main Line Between Stamford and Devon, Six Cars

Test No.	Time From	Time To	Filters Per Car	Train From	Train To	Maximum Metallic Noise Measured at Bridgeport, on New Haven-Harlem River Line, Db No. 7 West	No. 7 East
1	9:10 a.m.	11:06 a.m.	0	Stamford	Devon	26.5	10
2	11:07 a.m.	12:03 p.m.	1	Devon	Stamford	25	10
3	12:04 p.m.	12:47 p.m.	2	Stamford	Bridgeport	20.5	
4	2:27 p.m.	2:58 p.m.	0	Bridgeport	Stamford	28.5	

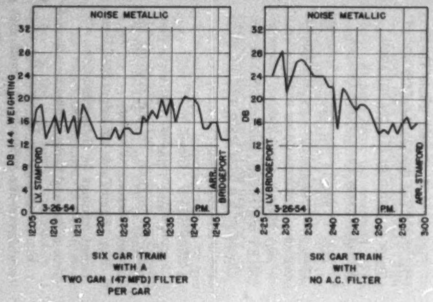


Fig. 17. Data obtained on railroad test circuit with rectifier cars on the main line between Stamford and Bridgeport

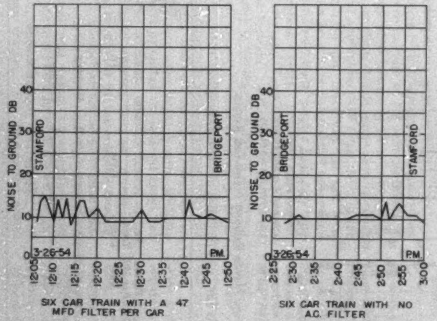


Fig. 18. Data obtained on SNET test circuit with rectifier cars on the main line between Stamford and Bridgeport

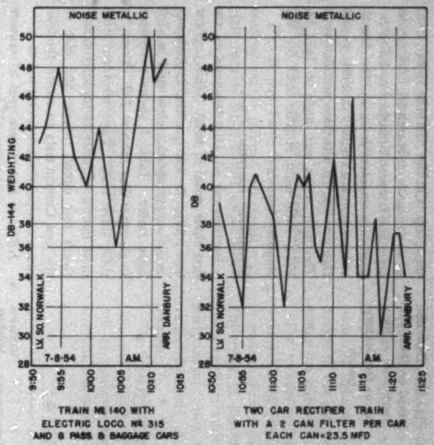


Fig. 19. Data obtained on railroad test circuits from tests on the Danbury branch

Table VIII. SNET Noise-Test Results

Test Runs on Main Line Between Stamford and Bridgeport With MU Cars, March 26, 1954

Test No.	Time		No. of MU Cars	Filters Per Car	Train		Maximum Noise to Ground
	From	To			From	To	
1	12:04	12:47 p.m.	6	2	Stamford	Bridgeport	15
2	2:27	2:56 p.m.	6	0	Bridgeport	Stamford	14

Table IX. New Haven Railroad Noise-Test Results

Test Runs of MU Cars, July 8, 1954, on Danbury Branch

Test No.	Time		No. of MU Cars	Filters Per Car	Train		Maximum Metallic Noise, Phantom Circuit, Db
	From	To			From	To	
1	10:51	11:22 a.m.	2	2	South Norwalk	Danbury	46
2	12:33	1:21 p.m.	2	2	Danbury	South Norwalk	46
3	2:38	3:17 p.m.	2	2	South Norwalk	Danbury	48
4	9:51	10:12 a.m.	train no. 140, electric locomotive no. 315		South Norwalk	Danbury	50

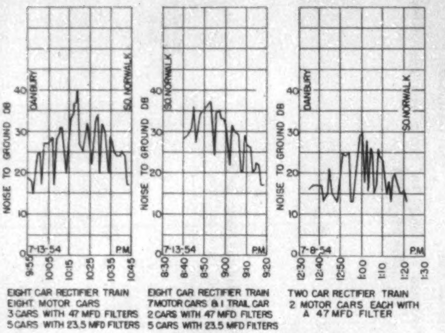


Fig. 20. Data obtained on SNET test circuit with rectifier cars on the Danbury branch

Table X. SNET Noise-Test Results

Test Runs on Danbury Branch With MU Cars, July 8 and 13, 1954

Test No.	Time		No. of MU Cars	Filters Per Car	Train		Maximum Noise to Ground
	From	To			From	To	
1	9:56	10:43 p.m.	8 { 3 5	2 1	Danbury	South Norwalk	40
2	8:31	9:16 p.m.	8 { 3 5	2 1	South Norwalk	Danbury	37.5
3	12:33	1:21 p.m.	2	2	Danbury	South Norwalk	30

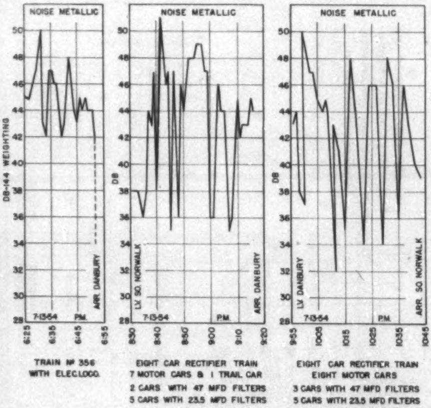


Fig. 21. Data obtained on railroad test circuit from tests on the Danbury branch

Table XI. New Haven Railroad Noise-Test Results

Test Runs of MU Car, July 13, 1954, on Danbury Branch

Test No.	Time		No. of MU Cars	Filters Per MU Car	No. Train Cars	Train		Maximum Metallic Noise, Phantom Circuit, Db
	From	To				From	To	
1	6:24	6:52 p.m.	train 356 and electric locomotive			South Norwalk	Danbury	50
2	8:31	9:16 p.m.	7 { 2 5	2 1	1	South Norwalk	Danbury	51
3	9:56	10:43 p.m.	8 { 3 5	2 1		Danbury	South Norwalk	50

the New Canaan branch. The third series was made on March 26, 1954, on the main line between Stamford and Bridgeport with six rectifier cars. The fourth series was made on July 7, 1954, on the Danbury branch with two rectifier cars. The fifth series was made on July 13, 1954, with eight rectifier cars on the Danbury branch.

MEASURING INSTRUMENTS

Rectifier car no. 4400 was used on the first three series of tests. This test car was equipped with meters, cathode-ray oscilloscopes, speedometer, and a *2-B* noise meter. Table I shows maximum single-car-trolley ampere $I \cdot T$ values for these three tests. Tape recorders were used during all tests to record the railroad's test circuit noises; and also in the New Canaan tests to record the telephone company's test circuit noises.

2-B noise meters were used to measure metallic noise on all railroad test circuits. Noise to ground could not be measured due to drainage circuit connections.

2-B noise meters were used to obtain noise to ground readings on selected SNET telephone circuits. Metallic noise was too low to measure. All db (decibel) readings taken were with 144 receiver weighting.

DISCUSSION OF TEST DATA

First Series of Tests at New Canaan

Fig. 9 and Table II show the noise readings obtained on the railroad test telephone line at New Canaan with single-phase series-motor motive power.

Fig. 10 and Table III show the corresponding reading for one and two rectifier MU cars with and without a-c filters.

Figs. 11 and 12 and Table IV show the SNET test line readings, for both a-c motor and rectifier MU car operation.

Second Series of Tests at New Canaan

Figs. 13 and 14 and Table V show the noise readings obtained on the railroad test circuit with one to six rectifier cars.

Figs. 15 and 16 and Table VI show the SNET test circuit readings for these conditions with a 6-car train with a-c series-motor locomotive.

Third Series of Tests, on Main Line

Fig. 17 and Table VII show the noise readings obtained on the railroad test circuit with six rectifier cars.

Fig. 18 and Table VIII show SNET test circuit readings for these conditions.

Fourth Series of Tests, on Danbury Branch

Fig. 19 and Table IX show the noise readings on the railroad test circuit with two rectifier cars and with a single-phase series-motor locomotive.

Fig. 20 and Table X show SNET test circuit reading under these conditions except no tests were made during a-c series-motor operation.

Fifth Series of Tests, on Danbury Branch

Fig. 21 and Table XI show the noise readings on the railroad test circuit with eight rectifier cars and with an a-c series-motor locomotive.

Fig. 20 and Table X show the corresponding SNET test circuit readings except no tests were made during a-c series-motor operation.

NOISE, MU RECTIFIER MOTIVE POWER

The most predominant and most characteristic sound produced by nonfiltered single-phase rectifier motive power is a "buzz" similar to the sound of a small buzzer. The buzz in the communication circuit is induced by the high rate of change in the catenary current and flux during the angle of overlap.

The voltage induced in the exposed communication circuit appears as a pulse in voltage during the angle of overlap. The value of induced voltage for any given exposure varies directly with the load current being commutated and it varies inversely with the angle of overlap.

The pulse in voltage in the communication circuit occurs every half-cycle and will be positive and negative during alternate half-cycles.

Conclusions

The test results indicate the following:

1. A large number of rectifier motive-power units do not have a directly additive effect on each other.
2. MU rectifier motive power without a-c filters produces approximately the same noise levels as existing motive power under similar load conditions.
3. MU rectifier motive power with a-c filters produces lower noise levels than existing motive power under similar load conditions.

Reference

1. RECTIFIER MOTIVE POWER—INDUCTIVE COORDINATION CONSIDERATIONS, E. B. King, E. H. Gordon, L. J. Hibbard. *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 107–18.

◆

Discussion

E. A. Binney (English Electric Company, Ltd., Bradford, England): In this very interesting paper the authors have shown records which indicate the noise found in open telephone lines paralleling a track, under varying current conditions in the catenary. It is, however, difficult to compare the results obtained with the locomotive-hauled trains and the MU cars with rectifiers, since the corresponding current are not given. It is clear that the MU cars cause no direct additive effect, but would this be true of a locomotive having, say, the power of six MU cars?

Two years ago a 25-cycle 6,600-volt system of the British Railways was converted to a 50-cycle 6,600-volt system, and the original MU cars having two single-phase motors were replaced by new cars with rectifier equipment, in this case there being four motors per car of 200 hp.

Prior to the change, noise tests were carried out on a telephone circuit paralleling the tracks, using a single motor coach (with two 200-hp single-phase motors) taking about 40 amp from the overhead. The psophometer readings obtained reached values of 8 millivolts, this being nearly twice the recognized maximum permissible in the United Kingdom, although the telephone lines had been used under these conditions for many years (under protest). Subsequently, similar tests were carried out with a single rectifier-type MU car drawing a maximum current of 200 amp during starting. The corresponding readings reached a value of 30 millivolts. A further test at 60 line amp gave 11 millivolts. From these results it appears that the interference value is roughly proportional to the current drawn by the rectifier car, and that at equal currents for the 25-cycle single-phase load and rectified 50-cycle load the interference is of the same order.

As might be surmised, these telephone wires were cabled and buried before operation on the new system commenced. I think it can be assumed that this step will be necessary on all 50/60-cycle single-phase traction systems, whatever type of motive power is employed.

Of greater interest, in my opinion, is the possible effect of harmonics due to 50-cycle traction on the power system or systems supplying the substations. It would be interesting to know the magnitude of the various harmonics due to traction currents which might be measured on the secondary and primary side of the West Farm supply, which would appear to be the only station connected to the system having nontraction demand.

H. S. Ogden (General Electric Company Erie, Pa.): The paper provides a record of a number of miscellaneous tests with reference to the subject of telephone interference on a particular railroad. The data given do not lend themselves readily to other applications. They are valuable in that they provide a qualitative summary of what happens when rectifier motive-power units are applied in rather large numbers.

Having been invited to participate as an observer in some of the tests described, it was my good fortune to be able to obtain some readings which may have a more general application than those given in the paper. On the first tests on the New Canaan branch, it was possible to measure the total rail return current from the train undergoing tests between Springdale Cemetery and New Canaan Station at the former location. Although the scale of the ammeter in use was such that it would not indicate currents under 120 amp (approximately current for two cars), fairly complete observations were made on a 4-car train with zero, one, and two capacitor cans per car. A shunt and noise meter in series with the ammeter permitted the inherent noise, and thus the telephone influence factor, of the propulsion current to be determined. When multiplied by the ammeter reading, the $I \cdot T$ product for the circuit became available. At the same instant that these meter readings were taken, corresponding noise measurements were made on the freight and ticket office telephone line at the New Canaan Station. The information so obtained has been plotted in Fig. 22. The curves indicate a reasonably close correlation between the $I \cdot T$ product for the different numbers of capacitors per car and the noise reading on the phone line. Within a reasonable amount of error, considering the test conditions, the curves show that doubling the $I \cdot T$ product changes the db level on the phone line by 6 db, which, because of the logarithmic nature of the db unit, is the correct ratio. The following simple manipulation proves this statement. By definition

$$\mathrm{db} = 20 \log_{10}(I \cdot T) \qquad (1)$$

Doubling the $I \cdot T$ product gives

$$\mathrm{db}' = 20 \log_{10}(2 \times I \cdot T) \qquad (2)$$

$$\mathrm{db}' = 20 (\log_{10} 2 + \log_{10} I \cdot T) \qquad (3)$$

$$\mathrm{db}' = 20 \log_{10} 2 + 20 \log_{10} I \cdot T \qquad (4)$$

Substituting the numerical value of 20 $\log_{10}$ 2 and db from equation 1 and rearranging gives

$$\mathrm{db}' = \mathrm{db} + 6$$

Thus doubling the $I \cdot T$ product increases the original db level by six units.

The $I \cdot T$ product decrease from 7,800 for no cans to approximately 2,000 with two cans per car, and the telephone-line noise level decreased from 51.5 db to 44 db. Talking over the line was entirely possible with a noise level of 51.5 db resulting from the rectifier. The buzz of the rectifier seemed to fade out when voice frequencies were present and was only noticeable when there was no talking.

A 1955 paper[1] contains a discussion on some of the theoretical aspects of telephone interference, and included therein is an oscillograph wave shape (Fig. 9) showing an induced wave on a telephone line on the test floor. A wave shape identical in all respects to this was obtained on the freight and ticket office telephone line at the New Canaan Station during these tests. A copy of this oscillogram is shown in Fig. 23. A single capacitor per car sufficed to remove the spike shown on the oscillogram.

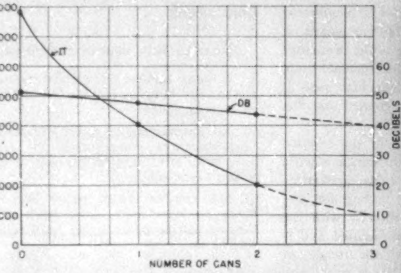


Fig. 22 (above). Tests of 4-car train on New Canaan branch, March 24, 1957; total I·T from Springdale Cemetery readings, and db metallic on freight and ticket office line versus number of cans per car

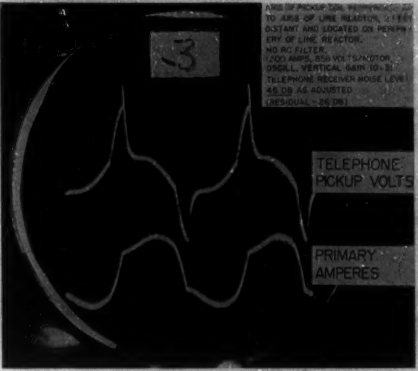


Fig. 23 (right). Oscillogram of test on ticket office telephone line at New Canaan branch

Where observations were made on telephone lines encased in cable, it was impossible to obtain the characteristic "pick-up" wave shape mentioned in the foregoing and, as stated in the paper, the influence was of very small proportions.

The authors omitted any reference to the use of tape recorders in studying telephone influence on the New Haven Railroad. An excellent presentation describing the use of tape recorders in connection with tests on the Pennsylvania Railroad was made by Mr. Hibbard and others (see reference 1 of the paper). I believe the use of tape recorders would bear mentioning again since, on some of the tests covered by the paper, tape recorders were placed at a number of locations on the New Haven telephone system, such as at the New Canaan ticket office, the Bridgeport wire chief's office, and at the New Haven dispatcher's office. These were in operation throughout the tests on the New Canaan branch and permitted personnel, then busily engaged with the conduct of the tests at the New Canaan ticket office, to study the record of test noise on the telephone lines at these specified points at their leisure a day or so later.

REFERENCE

1. CONSIDERATIONS IN THE DEVELOPMENT OF A HIGH-POWER RECTIFIER LOCOMOTIVE, H. S. Ogden. *AIEE Transactions*, vol. 74, pt. II, July 1955, pp. 169–76.

H. F. Brown (Gibbs & Hill, Inc., Pennsylvania Station, New York, N. Y.): Those familiar with inductive co-ordination problems involving single-phase railway-traction circuits and the paralleling of communication lines will recognize in this paper familiar territory and utilities involved. Ever since 1909, the New Haven Railroad and the SNET, with the American Telephone and Telegraph Company and Bell Laboratory eingineers, have closely co-operated in the understanding and solutions of their mutual problems due to induction. That both the railway and the communication systems have successfully extended their facilities in this territory during the intervening years indicates how satisfactory have been the relations between these companies, and how successful have been the mitigative methods developed.

The information presented will not only be of value when future equipment of this type is considered for use in the United States, but will be studied with keen interest abroad, where rectifier equipment is being used or considered for single-phase railway electrification using commercial frequency.

The harmonics from rectifier apparatus whether used in fixed substations for d-c traction, or on mobile equipment of the type described in the paper, have always presented a problem to the communication interests. It is gratifying to learn from the test data and experience already accumulated that this new type of motive power for railway electrification, which combines so well the best features of the d-c and a-c systems, requires but simple and relatively inexpensive mitigative apparatus. The railroad, the telephone, and manufacturing companies involved are all to be congratulated and complimented on another good example of engineering co-operation.

It seems fitting here to pay tribute to one of the authors of this paper, L. J. Hibbard. That the rectifier car and locomotive have been so successfully developed is due largely to his untiring efforts. The electrical industry and the railroads are greatly indebted to him for his enthusiasm in the application of this apparatus, and for his contributions to the solution of the engineering problems involved.

L. J. Hibbard, F. T. Garry, and G. N. Loomis: It is predicted that communication inductive co-ordination considerations in connection with single-phase rectifier motive-power applications to the world's railroads will become increasingly unimportant in the future.

From the communication standpoint, the extensive replacement by railway and telephone companies of open wire with cable, with its inherent shielding and close pair spacing, as well as the increasing use of carrier and short wave have made the communication plant less susceptive. Another factor has been the general policy of telephone companies for many years to avoid long and close inductive exposures to existing or potential railroad electrifications.

Our experiences in the United States to date indicate the only inductive communication problems that may be involved with applications of rectifier motive power to new or existing electrifications will be the railroads' own private communication lines. These problems can be met on existing communication lines where trouble is encountered by simple and inexpensive filtering equipment on the motive-power units.

These problems will disappear on the modern railroads of the future, equipped with modern communication systems. Hurricanes, ice storms, and other acts of nature combined with economics will force the modern railroad to use modern communication systems even if they are not electrified.

Extensive electrification is taking place in other parts of the world with rectifier motive power at commercial frequencies. To our present knowledge, no communication problems have been encountered and it has not been necessary to apply filtering equipment to the rectifier motive-power units.

In reply to Mr. Binney's question we did not measure the rectifier harmonics on the primary side of the West Farm substation. However, tests made in the laboratory at East Pittsburgh indicate that existing New Haven motive-power equipment with single-phase series commutator motors produce as high a value of harmonics in the supply system as are produced with the new rectifier motive units.

Several studies have been made in recent years relative to supplying power to single-phase railways from existing power systems at commercial frequency. The studies indicate that such a procedure is entirely practical. This conclusion is borne out by experiences in France.

Some Applications of Magnetic Amplifiers in Aircraft Generator Protective Systems

D. L PLETTE
ASSOCIATE MEMBER AIEE

J. W. BUTLER
ASSOCIATE MEMBER AIEE

CONSIDERABLE progress has been made during the years 1951 through 1956 toward improving the performance and reliability of aircraft generator regulation[1] and excitation systems. Much of this progress has been made possible by the use of static magnetic components and metallic dry disk rectifiers combined in such a way that the result is good reliability and accurate, predictable performance.

During this same period the increasing dependance of the airplane upon its electrical system has made it necessary to consider protective systems to prevent the destruction of critical load equipment due to a failure in any part of the electrical system. The majority of the protective devices utilized in these systems depends on the calibrated operation of relays. These relays are normally operated with coil voltages which are fairly critical. At rated system voltage and frequency, the coil voltage is either slightly lower than that necessary to pick up the relay or slightly higher than that at which the relay will drop out. This condition is brought about by the desirability of detecting frequency and voltage changes in the order of 5% to 15% from rated. As a result of this practice, these protective devices are sensitive to acceleration, shock, and vibration and their exact operating points are often rather unpredictable. Since the relays are generally not designed to withstand the vibration forces present in aircraft with critical coil voltages applied, vibration isolators are required for the equipment to prevent malfunctioning of these devices.

It is felt that the reliability and accuracy of the protective system can be much improved by the use of magnetic amplifiers, which can be used in such a manner that they supply essentially full voltage or zero voltage to the coil of a relay, thereby making the relay relatively insensitive to acceleration, shock, and vibration. This paper will describe some applications of magnetic amplifiers in aircraft generator protective systems.

Reasons for Use of Magnetic Amplifiers

Magnetic amplifiers provide an accurate and predictable means of obtaining power amplification. It is therefore possible to operate voltage comparison circuits and frequency comparison circuits at a fairly low power level. The error signal obtained from these comparison circuits can then be amplified to a power level sufficient to operate a reliable standard relay, supplying essentially full-on or full-off voltage to the relay coil.

The resulting use of standard, noncritical relays enables the protective system to withstand rigid shock, vibration, and acceleration specifications without the need for shock mounts or vibration isolators. Magnetic amplifiers also provide an accurate and reliable means of producing time delay. By utilizing the magnetic properties of the core, fixed time delay as well as inverse time delay can be obtained. By the proper selection of core materials and rectifiers, magnetic amplifiers will perform accurately and reliably over the range of −55 to +125 C (degrees centigrade).

Circuit Description

MAGNETIC-AMPIFIER UNDERVOLTAGE RELAY

Individual phase undervoltage protection is often required in protective systems to protect load equipment from sustained undervoltage because of any of the following:

1. Loss of excitation.
2. Open-generator phase windings or feeder (under conditions of system loading which produce undervoltage on one or more phases).
3. Bus faults or feeder faults which are not cleared either by protective devices or by burning clear.

Individual phase undervoltage protection may be obtained by utilizing a magnetic amplifier as shown in Fig. 1. The voltage on each phase (E_{ac}) is rectified and filtered to produce voltage (KE_{dc}) which is compared to the reference voltage. With rated voltage on each phase, the direct voltage produced by each sensing network is higher than the reference voltage and rectifiers 1REC, 2REC, and 3REC are blocking. If any one or all three direct voltages drop below the reference voltage, current flows from the reference through the control winding and rectifier. This causes the magnetic amplifier to reduce its output to a very low value, thereby de-energizing the relay. As soon as all three direct voltages are equal to or higher than the reference voltage, the magnetic amplifier will increase its output and energize the relay. Different settings of the undervoltage relay are obtained by the use of different fixed resistors for 1*R*, 2*R*, and 3*R*.

MAGNETIC-AMPLIFIER OVERVOLTAGE RELAY

Highest-phase overvoltage protection is commonly required in protective systems to give the maximum overvoltage protection to load equipment. This protection is characterized by an inverse time-voltage relationship. Extreme overvoltages are allowed to exist for only a very brief time. On the other hand, slight overvoltages are allowed to exist for a longer time before the overvoltage protection is energized.

Highest-phase overvoltage protection with inverse time delay may be obtained by utilizing a magnetic amplifier as shown in Fig. 2.

The three line-to-neutral voltages E_{Aac}, E_{Bac}, E_{Cac} are rectified and filtered in such a manner that the output voltage E_{0dc} is roughly proportional to the highest of the three line-to-neutral voltages. A portion of this voltage is then compared with the voltage reference. Under rated voltage conditions, this voltage $E_{2R} = [2R/(1R+2R)]\times E_{0dc}$ is less than the reference voltage and rectifier 1REC is blocking. As the voltage on any one or all phases of the system rises, the voltage E_{2R} will finally exceed the reference voltage and cause voltage to be impressed across control winding F_1-F_2 on the magnetic amplifier (1*SX*). The characteristic of the voltage across this winding as a

---

Paper 57-481, recommended by the AIEE Air Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE East Central and Middle Eastern District Meeting and Air Transportation Conference, Dayton, Ohio, May 7–9, 1957. Manuscript submitted February 7, 1957; made available for printing July 29, 1957.

D. L. PLETTE and J. W. BUTLER are with the General Electric Company, Waynesboro, Va.

The authors wish to acknowledge the assistance of H. H. Britten, P. D. Corey, and J. L. Lotts, who contributed substantially to the development work on which this paper is based.

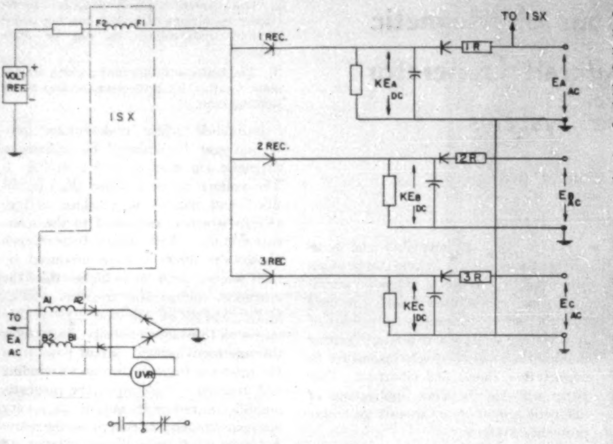


Fig. 1. Undervoltage relay circuit

function of line-to-neutral voltage is shown in Fig. 3. The exact value of this voltage for a particular value of E_{0dc} will depend upon the reference voltage and the values of resistors $1R$, $2R$, and $3R$.

An understanding of how the magnetic amplifier produces a time delay characteristic may be obtained by considering Fig. 4, which shows a hysteresis loop of the square-loop core material utilized in the magnetic amplifier. By means of bias winding F_3-F_4, the core material is initially saturated in the negative direction. By winding gate windings A_1-A_2 and B_1-B_2 with a very large number of turns, the a-c flux excursion in the core can be made very small compared to the saturation flux density of the material. Its effects can therefore be neglected in calculating the flux change caused by application of the voltage E_c to the control winding F_1-F_2 of the magnetic amplifier.

The time delay characteristics of this magnetic amplifier thus depend on the voltage E_c as follows:

$$E=N\frac{d\phi}{dt}\times 10^{-8} \tag{1}$$

$$E_c=N_cA_{Fe}\frac{dB}{dt}\times 10^{-8} \tag{2}$$

$$\int_0^t E_c dt=\frac{N_cA_{Fe}}{10^8}\int_{-B_s}^{+B_s} dB \tag{3}$$

$$K_{volt-sec}=\int_0^t E_c dt=\frac{N_cA_{Fe}2B_s}{10^8} \tag{4}$$

These equations indicate that for a given magnetic amplifier with saturation flux density B_s, control winding turns N_c, and total cross sectional area of both cores equal to A_{Fe}, a definite number of volt-seconds are required to cause the flux density to change from $-B_s$ to $+B_s$ or a total excursion of $2B_s$. At such time as the cores are saturated, the output voltage of the magnetic amplifier will suddenly increase causing the relay to be energized.

When the overvoltage is removed, thereby reducing the voltage E_c to zero, the output of the magnetic amplifier is quickly reduced to a very low value, thus de-energizing the relay. The bias winding F_3-F_4 will then cause the cores to be reset to the negative saturation value. Complete resetting is normally accomplished in 1 to 2 seconds after which the time delay action can be repeated with excellent accuracy.

$$K_{volt-sec}=\int_0^t (E_{ac}-E_{ut})dt \tag{5}$$

where $E_{ac}>E_{ut}$

Typical values of K required for use with present-day excitation systems vary from 2.0 to 5.0 volt-seconds.

The ultimate trip voltage of the relay is determined primarily by the reference voltage, the ratio of E_{ac} to E_{0dc}, and the resistance values of $1R$ and $2R$. The volt-second characteristic is then determined primarily by the core characteristics, the number of control winding turns, and the resistance of $3R$. A family of constant volt-second curves is shown in Fig. 5.

Once a particular volt-second characteristic has been established for the overvoltage relay, it is possible by analysis of oscillographic data to determine whether a particular overvoltage transient in the system would cause tripping of the overvoltage relay. If not, the amount of margin remaining before tripping occurs can be easily determined.

MAGNETIC-AMPLIFIER FIXED TIME DELAY RELAY

In the magnetic-amplifier overvoltage relay, the inverse time delay characteristic is obtained by applying varying voltages to the control winding F_1-F_2 of the amplifier. If, instead, a fixed control winding voltage is applied, the time required

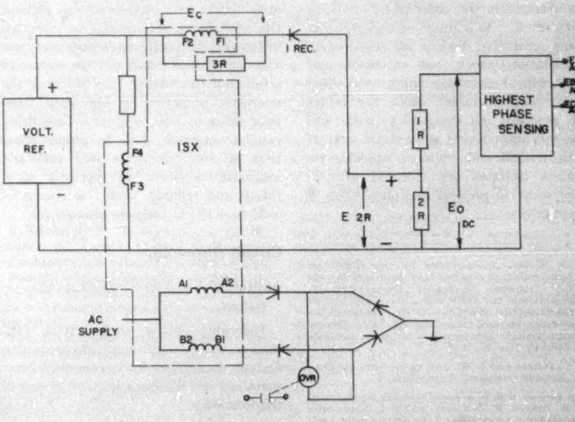


Fig. 2. Overvoltage relay circuit

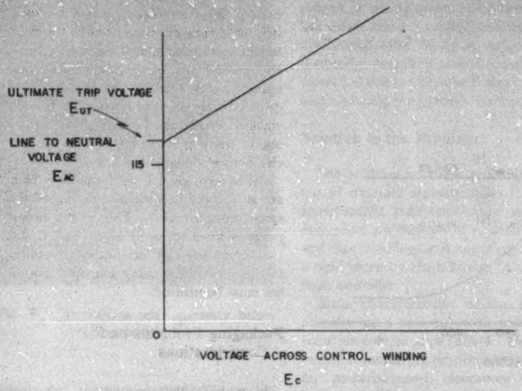


Fig. 3. Overvoltage relay comparison circuit characteristic

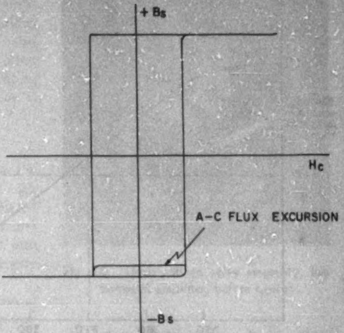
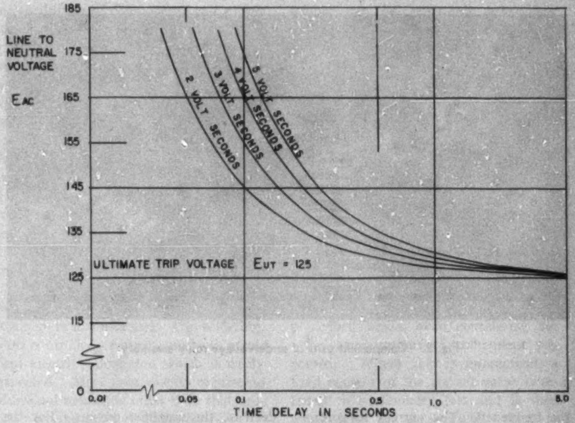


Fig. 4. Hysteresis loop of square loop core material

Fig. 5 (below). Overvoltage relay time delay curves



to change the flux level in the cores from negative to positive saturation will be a fixed time. Thus the same basic magnetic-amplifier time delay circuit can be utilized to obtain a fixed time delay as well as an inverse time delay.

Magnetic-Amplifier Underfrequency Relay

Underfrequency protection is often required in aircraft electric systems to prevent overheating of various types of load equipment. Transformers, magnetic amplifiers, and induction motors are examples of loads which cannot normally tolerate application of rated voltage at frequencies under their minimum design frequency.

An underfrequency may be brought about by any number of factors, such as:

1. Engine idle at low speed on direct-driven generators.
2. Failure of constant speed drive or associated controls.
3. Excessive load torque requirement imposed on a constant speed drive during a low-input speed or low-input energy condition.
4. Jet engine failure and windmilling during flight (even with constant speed drive).
5. Short-time underfrequency during normal engine shutdown and starting.

One method of providing underfrequency protection is the use of a magnetic-amplifier type of underfrequency relay. The circuit for a magnetic-amplifier type of underfrequency relay is shown in Fig. 6. The current in control winding F_3-F_4 of the magnetic amplifier is proportional to the voltage applied to the circuit and is essentially independent of

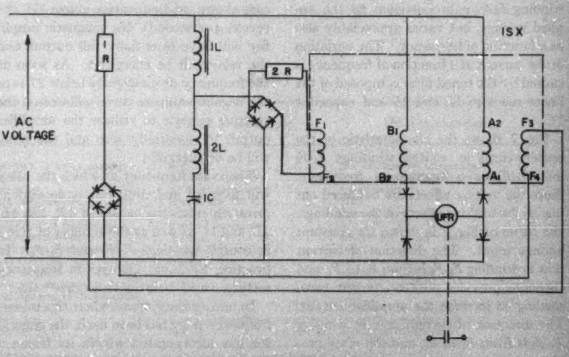


Fig. 6. Underfrequency relay circuit

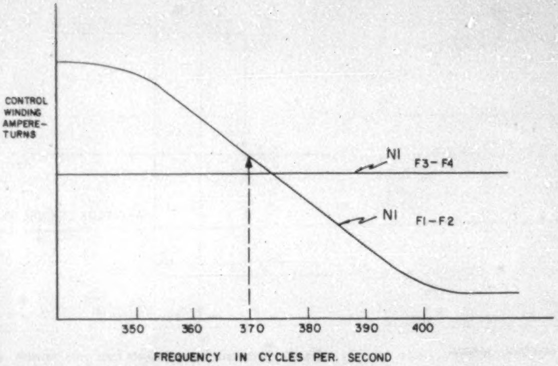


Fig. 7. Characteristics of frequency comparison circuit with constant input voltage

Fig. 8. Component parts of undervoltage relay assembly

the frequency. The current in control winding F_1-F_2 is proportional to the applied voltage, but varies appreciably also as a function of frequency. The variation of the current as a function of frequency is caused by the tuned filter composed of the linear reactors $1L$ and $2L$ and capacitor $1C$.

Fig. 7 shows the characteristic of the ampere-turns in control windings F_1-F_2 and F_3-F_4 as a function of frequency. Since the voltage effects are balanced out due to the bucking effects of the windings, the curve of Fig. 7 is shown for constant voltage input. The direction of current flow in winding F_3-F_4 is from F_3 to F_4, and therefore provides positive ampere-turns tending to increase the amplifier output. The direction of current flow in winding F_1-F_2 is from F_2 to F_1, and therefore provides negative ampere-turns tending to decrease the amplifier output. For the case shown, at frequencies above 370 cps (cycles per second), the magnetic amplifier will have essentially full output and the relay will be energized. As soon as the frequency drops slightly below 370 cps the negative ampere-turns will exceed the positive enough to reduce the amplifier output to essentially zero and the relay will be de-energized.

The exact frequency at which the relay will pick up and drop out is dependent primarily upon the values of $1R$, $2R$, $1L$, $2L$, and $1C$ as well as the number of turns in control windings F_1-F_2 and F_3-F_4. In practice, for slight changes in frequency setting, $1R$ is varied in resistance value.

In most of the systems where this underfrequency relay has been used, the generator has incorporated within its frame a small permanent-magnet auxiliary generator. This auxiliary generator is utilized for a number of purposes in the regulation, excitation, and protection systems. By sensing the frequency of this permanent-magnet generator, the underfrequency relay can provide performance which is unaffected by loss of voltage on one or more phases of the main aircraft electrical system.

An inverse time underfrequency characteristic may be obtained by use of the same basic circuit.[2] The tuned circuit range of linearity must be extended, and to accomplish this an amplifier similar to the overvoltage relay magnetic amplifier must be utilized.

Packaging Problems and Considerations

In making the product or physical design of a protective control panel utilizing magnetic amplifier circuits, a number of considerations confront the product designer. The following is only a partial list of items for consideration:

1. General specifications which equipment must meet.
2. Adaption of the product design to the available components that individually will meet the electrical and environmental requirements.
3. Adaption of the design for ease of manufacturing, testing, and servicing.

GENERAL SPECIFICATION REQUIREMENTS

The unit must operate reliably under the normal military aircraft environmental conditions, including high and low temperature, altitude, humidity, salt spray, fungus, shock, acceleration, vibration, and others. Perhaps the requirements presenting the greatest problems to the designer are the high temperatures (120 degrees centigrade, minimum requirement), vibration (resonances resulting from an applied vibration of 10 *g* from 60 to 500 cps minimum), and shock (equivalent to 15 *g* as a minimum requirement). Minimum size and weight are obviously of the utmost importance.

As additional design objectives, it is considered preferable to avoid the use of any adjustable resistors or potentiometers, as well as vibration isolators. Adjustable resistors and potentiometers are potential contributors to unreliability in aircraft equipment. This is because of their basic construction requirement for a moving member, their exposed electrical contact surfaces, and their susceptibility to damage from well-meaning but inexperienced aircraft mechanics. Vibration isolators rob valuable space and result in

Fig. 9. Overvoltage relay assembly before casting, bottom view

larger and therefore potentially heavier products.

COMPONENT CONSIDERATIONS

Not only have protective systems gradually included more protective and control functions, but the application of magnetic amplifiers to these individual functions results in the use of larger numbers of components in each function. For example, the undervoltage relay circuit described in this paper actually consists of 29 electrical components, including rectifiers, resistors, capacitors, a relay, and a saturable reactor; see Fig. 8. Admittedly, most of these components are extremely small and dissipate only negligible amounts of power. However, since failure of any component could result in unsatisfactory operation of the undervoltage relay, considerable thought must be given to the mounting and the wiring of each component. If conventional packaging techniques are used, an extreme amount of attention and time must be given to detailed design, quality of workmanship, and thorough inspection, in order to achieve the required reliability. In spite of the fact that each individual component can pass rigid vibration specifications, it may be difficult for the over-all assembly to meet these requirements, without the use of vibration isolators.

ADAPTABILITY TO MANUFACTURING AND SERVICING

Products with large numbers of components tightly packaged are inherently difficult to manufacture. Only with extreme care on the part of the workman and the designer can a reliable assembly be built. In addition, if a unit fails to pass its electrical production test, it is often difficult to locate the faulty component or wiring error. The same problem exists when a malfunctioning unit is overhauled after being in service. Undoubtedly complete assemblies are often discarded due to the difficulties in locating and replacing the defective component.

Solution to the Problem

One solution to these problems is the use of modular subassemblies. In each subassembly, components are assembled and wired by the use of printed circuits and dip soldering, and are then cast into a rigid block by the addition of an epoxy resin and filler.

Some characteristics of the material actually used and an ideal casting material are compared in Table I. Due to the excellent insulating properties of the casting material used, clearances between points of different potential can be reduced without fear of failure. Failures due to shifting of components and wiring early in life are eliminated. All components and wiring within the cast block are protected from the adverse affects of humidity, salt spray, fungus, and altitude. Vibration problems are lessened since all components and wiring are held rigidly in place and vibrate as one mass. This prevents failure due to minor resonances at extremely high frequencies. Even short-time exposure to temperatures over the maximum continuous ambient temperature will not harm components since the heat actually takes appreciable time to reach the components.

The use of printed wiring reduces the possibilities of wiring errors and wire breakage in manufacture. Dip soldering, when controlled properly, produces a uniform electric connection which is easily inspected. Most components require no additional mounting other than that provided by the printed wiring board.

Fig. 10. Undervoltage relay assembly, less magnetic amplifier, before casting

A complete protective panel thus consists of a number of functional subassemblies plus a few miscellaneous components that do not lend themselves to casting due to their large power dissipation or large physical size. Each subassembly can be separately assembled, wired, tested, cast, and retested, prior to being placed in the final over-all equipment enclosure. Final-production electrical testing and inspection serve to check miscellaneous components and interconnecting wiring, as well as to recheck each function after its integration into the protective system. The use of the same functional subassemblies in different protective systems improves manufacturing efficiency and reduces product design time on new protective systems.

The use of adjustable resistors, rheostats, and potentiometers is eliminated wherever possible by designing circuits which do not require adjustment to meet prescribed limits, even considering the worst combinations of component tolerances. Where this is impractical, a fixed resistor of the correct value is selected in preliminary test and is then wired into the subassembly prior to casting. After casting, the subassembly

Table I

	Ideal Requirements	Actual Characteristics
Viscosity, prior to casting	low	low at elevated temperature
Pot life	as long as possible	29 days when stored at room ambient
Cure time	as short as possible	2 to 3 hours
Cure temperature	low, below 120 C	110 C
Shrinkage	kept to minimum	approximately 1%
Coefficient of expansion	kept to minimum	3.2×10^{-5}/C
Heat conductivity	good	less than 4.8×10^{-4} cal/sec/cm^2/C/cm.*
Exothermic temperature	as low as possible	160 C
Bond strength	high	excellent to metals
Shear strength	high	excellent to 160 C
Dielectric strength	high	over 400 volts/mil
Resistance to environmental conditions	high	excellent
Insulation	above 10 megohms per mil	over 100 megohms/mil
Density	low	0.036 pound per inch^3

* Calories per second per square centimeter per C per centimeter.

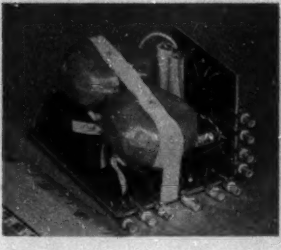


Fig. 11. Undervoltage relay assembly before casting

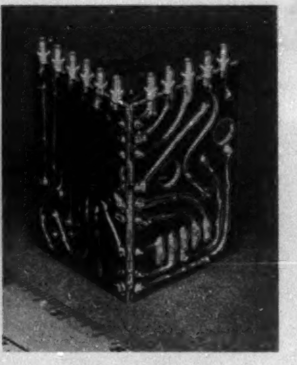


Fig. 12. Overvoltage relay assembly before casting, back view

will be interchangeable with all other similar cast blocks even though different resistors might have been selected for each block.

Testing and overhauling equipment in the field is greatly simplified by the cast subassembly approach. A brief check of the over-all equipment will generally indicate trouble in a particular function. By merely replacing that complete function, the equipment can be restored to normal operating condition without need for any adjustment. The cost of replacing complete subassemblies instead of individual components is at least partially offset by a simplified stocking problem and greater speed in repairing equipment.

PACKAGING OF SPECIFIC MODULAR BLOCKS

In the product designs of each of the modular blocks, there has been an attempt to secure mechanically all the components directly to the printed board in order to facilitate the assembling and to prevent wire breakage during manufacture. For example, the overvoltage relay shown in Fig. 9 has all its components mounted directly to the printed boards. In the undervoltage relay, the large number of parts mentioned previously (Fig. 8) are mounted directly to the two printed circuit boards with the exception of the saturable reactor; see Fig. 10. The two cores of the saturable reactor are separate for better space utilization and are held in proper position for casting by a short length of high-temperature glass tape; see Fig 11.

Fig. 13. Cast unit showing inserts for mounting

Since the height of the components varies greatly, mounting them on a single printed circuit board results in poor space utilization and extra weight in casting material. Using two printed circuit boards in parallel presents problems in assembly and test due to the relative inaccessibility of the components. The use of two printed boards mounted at right angles to each other affords a good compromise between ease of manufacture and minimum weight and size; see Figs. 9 through 11. The right-angle connections are made with wire jumpers.

Delicate components such as magnetic amplifiers which are sensitive to minute changes in pressure must be protected with a rubberlike compound before casting to prevent any changes in electrical characteristics caused by a difference in the thermal expansion of the component and the casting compound; see Fig. 11. The terminals used are of the solder type rather than the plug-in type. Better reliability is obtained from a soldered connection since the electrical contact is not dependent upon mechanical pressure and proper alignment. The terminals are placed on the edge of each printed board rather than on the surface of one board to make the casting operation easier; see Fig. 12. The assembly is placed in the mold with the terminals at the top, the casting compound is poured until the printed circuit board is covered, and only the terminals are left exposed. Steel inserts molded into each subassembly serve as a method of mounting the modular blocks; see Fig. 13.

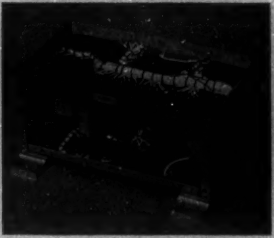


Fig. 14. Typical single-generator protective panel, less cover

Fig. 15. Single-generator protective panel

TYPICAL PROTECTIVE PANEL

A typical aircraft protective panel utilizing the modular construction is shown in Fig. 14. This unit provides highest-phase overvoltage protection, individual-phase undervoltage protection, underfrequency protection, transformer-rectification, and field flashing. The four cast assemblies are placed side by side in the rear of the panel with a sheet metal framework wrapped around them and a center support section in front. This enables all the assemblies to be securely fastened on at least three surfaces. This basic construction is extremely rigid and passes the vibration requirements easily.

The uncast components are mounted on

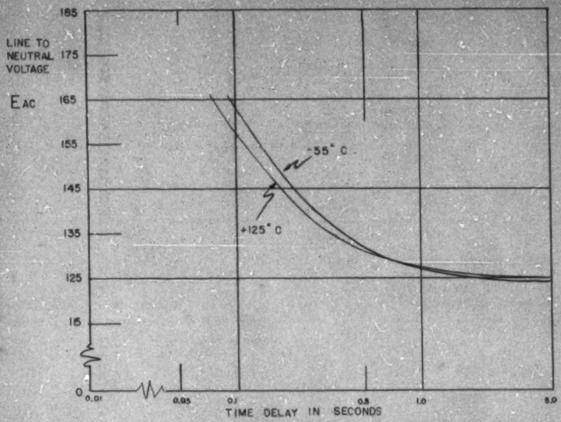


Fig. 16. Overvoltage relay variation with temperature

either the printed board or any of the available sheet metal surfaces. The complete protective panel weighing $6^1/_2$ pounds is shown in Fig. 15.

Results

1. The magnetic-amplifier type of overvoltage relay in its final form weighs 0.7 pound and is 2 by 2 by 2.8 inches. The electrical performance over the temperature range of −55 C to +125 C is shown in Fig. 16.

2. The magnetic-amplifier type of undervoltage relay weighs 0.75 pound and is 1.9 by 2.4 by 2.9 inches. Over the temperature range of −55 C to +125 C, the drop-out voltage varies ±2%. The differential between pickup and dropout is one volt, based on line-to-neutral voltage.

3. The magnetic-amplifier type of underfrequency relay weighs 0.95 pound and is 2.1 by 2.4 by 2.9 inches. Over the temperature range of −55 C to +125 C, the drop-out frequency varies ±3 cps. The differential between the pickup and dropout frequencies is 2 cps.

Conclusions

The use of magnetic amplifiers in aircraft generator protective systems results in highly accurate and predictable operation of the various protective functions. By eliminating the use of less reliable components such as adjustable resistors and poorly applied relays, the level of component reliability can be made very high. The use of cast subassemblies makes possible a high degree of reliability for the over-all assembly, without the requirement for vibration isolators. The use of modular subassemblies reduces product design time and simplifies the manufacturing and servicing of equipment. The size and weight of these protective devices are small, thereby making all the other advantages possible without increasing the weight and space requirements in the aircraft.

References

1. The Development of a Static Voltage Regulator for Aircraft A-C Generators, H. H. Britten, D. L. Plette. *AIEE Transactions*, vol. 74, pt. II, 1955(Jan. 1956 section), pp. 438–43.

2. Inverse-Time Underfrequency Relay, R. G. Hoft. *Ibid.*, vol. 75, pt. I, 1956 (Jan. 1957 section) pp. 625–29.

LATE DISCUSSION

The following discussion, based on a 1956 paper, was recently received from abroad.

Fundamental Relation of A-C Servo Systems

Discussion of paper 56-792, "Envelope Transfer Function Analysis in A-C Servo Systems," by M. Panzer, published in AIEE Transactions, part II, November 1956, pages 274–79.

Amos Nathan (Technion, Israel Institute of Technology, Haifa, Israel): What may well be called the fundamental relation of a-c servo systems was recently announced by Panzer and also by Levenstein.[1] Because of its importance, a much simplified proof is here presented.

Following Panzer's notation we state the problem as follows:

A linear transducer of transmittance $G(p)$, i.e.,

$$E_0(p)=G(p)E_i(p) \tag{1}$$

is fed with the input

$$e_i(t)=e_s(t)\cos\omega_c t=Re\{e_s(t)e^{i\omega_c t}\} \tag{2}$$

It is required to obtain the output $e_0(t)$ in the form

$$e_0(t)=Re\{[s_o(t)+iS_Q(t)]e^{i(\omega_c t+\phi)}\} \tag{3}$$

ϕ being an arbitrary constant.

To solve, let us transform equations 2 and 3, obtaining

$$E_i(p)=Re\{E_s(p-i\omega_c)\} \tag{2A}$$

$$E_0(p)=Re\{[S_D(p-i\omega_c)+ iS_Q(p-i\omega_c)]e\}^{i\phi} \tag{3A}$$

where we have used the well-known relation

$$\mathfrak{L}\{f(t)e^{\alpha t}\}=F(p-\alpha)$$

Substituting equations 2(A) and 3(A) into equation 1 and noting the reality of $G(p)$

$$Re\{[S_D(p-i\omega_c)+iS_Q(p-i\omega_c)]e^{i\phi}\} = Re\{G(p)E_s(p-i\omega_c)\} \tag{4}$$

which is satisfied by

$$S_D(p)+iS_Q(p)=G(p+i\omega_c)E_s(p)e^{-i\phi} \tag{5}$$

This solves the problem, $e_0(t)$ being given by equations 3 and 5.

Introducing

$$\begin{aligned} G_D(p)&=S_D(p)/E_s(p)\\ G_Q(p)&=S_Q(p)/E_s(p) \end{aligned} \tag{6}$$

and noting the reality of S_D and S_Q we can write equivalently:

$$\begin{aligned} G_D(p)&=Re\{G(p+i\omega_c)e^{-i\phi}\}\\ G_Q(p)&=Im\{G(p+i\omega_c)e^{-i\phi}\} \end{aligned} \tag{5A}$$

REFERENCE

1. ON THE DESIGN OF A-C NETWORKS FOR SERVO COMPENSATION, H. Levenstein. *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, New York, N. Y., Feb. 1957, pp. 39–55.

◆

ERRATA

"Design and Calculation of Induction-Heating Coils" by R. M. Baker, published in Applications and Industry, March 1957, pages 31–40.

Equation 25, page 34, should read:

$$X_e=\frac{8\pi^2 fN_c^2 10^{-9}}{R_e} \tag{25}$$

Equation 26, page 34, should read:

$$X_e=\frac{8\pi^2 fN_c^2 10^{-9}}{l_c}\left(\frac{l_c}{R_e}\right) \tag{26}$$